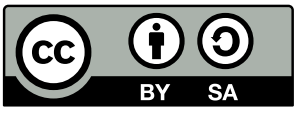
Inverse Problems

Lecture notes, Michaelmas term 2020
University of Cambridge

Yury Korolev and Jonas Latz

May 18, 2021



Contents

These lecture notes are based on the Inverse Problems course taught by Yury Korolev and Jonas Latz in Michaelmas term 2020 at the University of Cambridge.[1] Complementary material can be found in the following books, lecture notes and review papers:

1. Heinz Werner Engl, Martin Hanke, and Andreas Neubauer. *Regularization of Inverse Problems.* Springer, 1996.

2. Otmar Scherzer, Markus Grasmair, Harald Grossauer, Markus Haltmeier and Frank Lenzen. *Variational Methods in Imaging.* Springer, 2008.

3. Kristian Bredies and Dirk Lorenz. *Mathematical Image Processing.* Springer, 2018

4. Martin Benning and Martin Burger. *Modern regularization methods for inverse problems.* Acta Numerica, 27, 1-111 (2018)

   `https://www.cambridge.org/core/journals/acta-numerica/article/modern-regularization-methods-for-inverse-problems/1C84F0E91BF20EC36D8E846EF8CCB830`

5. K.Saxe. *Beginning Functional Analysis.* Springer, 2002

6. Masoumeh Dashti and Andrew M. Stuart, *The Bayesian approach to inverse problems*, Handbook of Uncertainty Quantification, 2016.

7. Jari Kaipio and Erkki Somersalo, *Statistical and computational inverse problems*, vol. 160 of Applied Mathematical Sciences, 2005.

8. O. Kallenberg, *Foundations of modern probability theory*, Springer, 1997.

9. Andrew M. Stuart, *Inverse problems: a Bayesian perspective*, Acta Numerica, 2010.

These lecture notes are under constant redevelopment and might contain typos or errors. We very much appreciate if you report any mistakes found (to y.korolev@damtp.cam.ac.uk or jl2160@cam.ac.uk). Thanks!

[1] `https://www.damtp.cam.ac.uk/research/cia/inverse-problems-michaelmas-2020`

Chapter 1

Introduction to Inverse Problems

Inverse problems arise from the need to gain information about an unknown object of interest from given indirect measurements. Inverse problems have several applications varying from medical imaging and industrial process monitoring to ozone layer tomography and modelling of financial markets. The common feature for inverse problems is the need to understand indirect measurements and to overcome extreme sensitivity to noise and modelling inaccuracies. In this course we employ both deterministic and probabilistic approach to inverse problems to find stable and meaningful solutions that allow us quantify how inaccuracies in the data or model affect the obtained estimate.

1.1 Well-posed and ill-posed problems

We start by considering the problem of finding $u \in \mathbb{R}^d$ that satisfies the equation

$$f = Au, \tag{1.1}$$

where $f \in \mathbb{R}^k$ is given. We refer to f as observed data or measurement and u as an unknown. The physical phenomena that relates the unknown and the measurement is modelled by a matrix $A \in \mathbb{R}^{k \times d}$. In real life the perfect data given in (1.1) is perturbed by noise and we observe measurements

$$f_n = Au + n, \tag{1.2}$$

where $n \in \mathbb{R}^k$ represents the observational noise.

We are interested in ill-posed inverse problems, where the inverse problem is more difficult to solve than the direct problem of finding f_n when u is given. To explain this we first need to introduce well-posedness as defined by Jacques Hadamard:

Definition 1.1.1. *A problem is called well-posed if*

1. *There exists at least one solution. (Existence)*
2. *There is at most one solution. (Uniqueness)*
3. *The solution depends continuously on data. (Stability)*

The direct or forward problem is assumed to be well-posed. The inverse problems are ill-posed and break at least one of the above conditions.

1. Assume that $d < k$ and $A : \mathbb{R}^d \to \mathcal{R}(A) \subsetneq \mathbb{R}^k$, where the range of A is a proper subset of $\mathbb{R}^k$. Furthermore, we assume that A has a unique inverse $A^{-1} : \mathcal{R}(A) \to \mathbb{R}^k$. Because of the noise in the measurement $f_n \notin \mathcal{R}(A)$ so that simply inverting A with the data given in (1.2) is not possible. Note that usually only the statistical properties of the noise n are known so we cannot just subtract it.

2. Assume next that $d > k$ and $A : \mathbb{R}^d \to \mathbb{R}^k$, in which case the system is underdetermined. We then have more unknowns than equations which means that there are several possible solutions.

3. Consider next case $d = k$ and there exist $A^{-1} : \mathbb{R}^k \to \mathbb{R}^k$ but the condition number $\kappa = \lambda_1/\lambda_k$, where λ_1 and λ_k are the biggest and smallest eigenvalues of A, is very large. Such a matrix is said to be ill-conditioned and is almost singular. In this case the problem is sensitive even to smallest errors in the measurement. Hence the naive reconstruction $\widetilde{u} = A^{-1}f_n = u + A^{-1}n$ does not produce a meaningful solution but will be dominated by $A^{-1}n$. Note that $\|A^{-1}n\|_2 \approx \|n\|_2/\lambda_k$ can be arbitrarily large.

The last part illustrates one of the key perspectives of inverse problem theory; How can we stabilise the reconstruction process while maintaining acceptable accuracy?

A deterministic way of achieving a unique and stable solution for the problem (1.2) is to use regularisation theory. In the classical Tikhonov regularisation a solution is attained by solving

$$\min_{u\in\mathbb{R}^d} \left(\|Au - f_n\|^2 + \alpha \|Lu\|^2 \right). \tag{1.3}$$

Above α acts as a tuning parameter balancing the effect of the data fidelity term $\|Au - f_n\|^2$ and the stabilising regularisation term $\|u\|^2$. The first half of the course will concentrate on regularisation theory.

Another way of tackling problems arising from ill-posedness is Bayesian inversion. The idea of statistical inversion methods is to rephrase the inverse problem as a question of statistical inference. We then consider problem

$$F = AU + N, \tag{1.4}$$

where the measurement, unknown and noise are now modelled as random variables. This approach allows us to model the noise through its statistical properties. We can also encode our *a priori* knowledge of the unknown in form of a probability distribution that assigns higher probability to those values of u we expect to see. The solution to (1.4) is so-called *posterior distribution*, which is the conditional probability distribution of u given a measurement m. This distribution can then be used to obtain estimates that are most likely in some sense. We will return to the Bayesian approach to inverse problems in the second half of the course

In this course we will concentrate on continuous inverse problems where in (1.1) and (1.2) $A : X \to Y$ is a linear or non-linear forward operator acting between some spaces X and Y, typically Hilbert or Banach spaces, the measured data $f_n \in Y$ is a function and $u \in X$ is the quantity we want to reconstruct from the data. Linear inverse problems include such important applications as computer tomography, magnetic resonance imaging and image deblurring in microscopy or astronomy. In other important applications, such as seismic imaging, the forward operator is non-linear (e.g., parameter identification problems for PDEs). Next we will take a look at some examples of linear and non-linear inverse problems to see what kind of challenges we face when trying to solve them.

1.2 Examples of inverse problems

1.2.1 Signal deblurring

The deblurring (or deconvolution) problem of recovering an input signal u form an observed signal

$$f_n(t) = \int_{-\infty}^{\infty} a(t-s)u(s)ds + n(t)$$

occurs in many imaging, and image- and signal processing applications. Here the function a is known as the blurring kernel.

The noiseless data is given by $f(t) = \int_{-\infty}^{\infty} a(t-s)u(s)ds$ and its Fourier transform is $\widehat{f}(\xi) = \int_{-\infty}^{\infty} e^{-i\xi t} f(t)dt$. The convolution theorem implies

$$\widehat{f}(\xi) = \widehat{a}(\xi)\widehat{u}(\xi),$$

and hence by inverse Fourier transform

$$u(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{it\xi}\frac{\widehat{f}(\xi)}{\widehat{a}(\xi)}d\xi.$$

However, we can only observe noisy measurements and hence we have on the frequency domain $\widehat{f_n}(\xi) = \widehat{a}(\xi)\widehat{u}(\xi) + \widehat{n}(\xi)$. The estimate u_{est} based on the convolution theorem is given by

$$u_{est}(t) = u(t) + \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{it\xi}\frac{\widehat{n}(\xi)}{\widehat{a}(\xi)}d\xi,$$

which is often not even well defined, since usually the kernel a decreases exponentially (or has compact support), making the denominator small, whereas the Fourier transform of the noise will be non-zero.

1.2.2 Heat equation

Next we study the problem of recovering the initial condition u of the heat equation from a noisy observation f_n of the solution at some time $T > 0$. We consider the heat equation on a torus $\mathbb{T}^d$, with Dirichlet boundary conditions

$$\begin{cases} \frac{dv}{dt} - \Delta v = 0 & \text{on } \mathbb{T}^d \times \mathbb{R}_+ \\ v(x,t) = 0 & \text{on } \partial\mathbb{T}^d \times \mathbb{R}_+ \\ v(x,T) = f(x) & \text{on } \mathbb{T}^d \\ v(x,0) = u(x) & \text{on } \mathbb{T}^d \end{cases}$$

where Δ denotes the Laplace operator and $\mathcal{D}(\Delta) = H_0^1(\mathbb{T}^d) \cap H^2(\mathbb{T}^d)$. Note that the operator $-\Delta$ is positive and self-adjoint on Hilbert space $\mathcal{H} = L^2(\mathbb{T}^d)$.

Given a function $u \in L^2(\mathbb{T}^d)$ we can decompose it as a Fourier series

$$u(x) = \sum_{n\in\mathbb{Z}^d} u_n e^{2\pi i\langle n,x\rangle},$$

where $u_n = \langle u, e^{2\pi i \langle n,x \rangle} \rangle$ are the Fourier coefficients, and the identity holds for almost every $x \in \mathbb{T}^d$. The L^2 norm of u is given by the Parseval's identity $\|u\|_{L^2}^2 = \sum |u_n|^2$. Remember that the Sobolev space $H^s(\mathbb{T}^d)$, $s \in \mathbb{N}$, consist of all $L^2(\mathbb{T}^d)$ integrable functions whose α^{th} order weak derivatives exist and are $L^2(\mathbb{T}^d)$ integrable for all $|\alpha| \leqslant s$. The fractional Sobolev space $H^s(\mathbb{T}^d)$ is given by the subspace of functions $u \in L^2(\mathbb{T}^d)$, such that

$$\|u\|_{H^s}^2 = \sum_{n \in \mathbb{Z}^d} (1 + 4\pi^2 |n|^2)^s |u_n|^2 < \infty. \tag{1.5}$$

Note that for a positive integer s, the above definition agrees with the definition given using the weak derivatives. For $s < 0$, we define $H^s(\mathbb{T}^d)$ via duality or as the closure of $L^2(\mathbb{T}^d)$ under the norm (1.5). The resulting spaces are separable for all $s \in \mathbb{R}$.

The eigenvectors of $-\Delta$ in $\mathbb{T}^d$ form the orthonormal basis of $L^2(\mathbb{T}^d)$ and the eigenvalues are given by $4\pi^2|n|^2$, $n \in \mathbb{Z}^d$. We can also work on real-valued functions where the eigenfunctions $\{\varphi_j\}_{j=1}^{\infty}$ comprise sine and cosine functions. The eigenvalues of $-\Delta$, when ordered on a one-dimensional lattice, then satisfy $\lambda_j \asymp j^{\frac{2}{d}}$. The notation $\asymp$ means that there exist constants $C_1, C_2 > 0$, such that $C_1 j^{\frac{2}{d}} \leqslant \lambda_j \leqslant C_2 j^{\frac{2}{d}}$.

The solution to the forward heat equation can be written as

$$v(t) = \sum_{j=1}^{\infty} u_j e^{-\lambda_j t} \varphi_j.$$

We notice that

$$\|v(t)\|_{H^s}^2 \asymp \sum_{j=1}^{\infty} j^{\frac{2s}{d}} e^{-2\lambda_j t} |u_j|^2 = t^{-s} \sum_{j=1}^{\infty} (\lambda_j t)^s e^{-2\lambda_j t} |u_j|^2 \leqslant C t^{-s} \sum_{j=1}^{\infty} |u_j|^2 = C t^{-s} \|u\|_{L^2}$$

which implies that $v(t) \in H^s(\mathbb{T}^d)$ for all $s > 0$.

We now have observation model

$$f_n = Au + n,$$

where $A = e^{T\Delta}$ and n is the observational noise. The noise is not usually smooth (the often assumed white noise is not even an L^2 function) and hence measurement f_n is not in the image space $\mathcal{D}(e^{T\Delta}) \subset \cap_{s>0} H^s(\mathbb{T}^d)$.

1.2.3 Differentiation

Consider the problems of evaluation the derivative of a function $f \in L^2[0, \pi/2]$. Let

$$Df = f',$$

where $D \colon L^2[0, \pi/2] \to L^2[0, \pi/2]$.

Proposition 1.2.1. *The operator D is unbounded from $L^2[0, \pi/2] \to L^2[0, \pi/2]$.*

Proof. Take a sequence $f_n(x) = \sin(nx)$, $n = 1, \ldots, \infty$. Clearly, $f_n \in L^2[0, \pi/2]$ for all n and $\|f_n\| = \sqrt{\frac{\pi}{4}}$. However, $Df_n(x) = n\cos(nx)$ and $\|Df_n\| = n \to \infty$ as $n \to \infty$. Therefore, D is unbounded. □

This shows that differentiation is ill-posed from L^2 to L^2. It does not mean that it can not be well-posed in other spaces. For instance, it is well-posed from H^1 (the Sobolev space of L^2 functions whose derivatives are also L^2) to L^2. Indeed, $\forall u \in H^1$ we get

$$\|Df\|_{L^2} = \|f'\|_{L^2} \leqslant \|f\|_{H^1} = \|f\|_{L^2} + \|f'\|_{L^2}.$$

However, since in practice we typically deal with functions corrupted by nonsmooth noise, the L^2 setting is practice-relevant, while the H^1 setting is not.

Differentiation can be written as an inverse problem for an integral equation. For instance, the derivative u of some function $f \in L^2[0,1]$ with $f(0) = 0$ satisfies

$$f(x) = \int_0^x u(t)\,dt,$$

which can be written as an operator equation $Au = f$ with $(A\cdot)(x) := \int_0^x \cdot(t)\,dt$.

1.2.4 Matrix inversion

In finite dimensions, the inverse problem (1.1) is a linear system. Linear systems are formally well-posed in the sense that the error in the solution is bounded by some constant times the error in the right-hand side, however, this constant depends on the condition number of the matrix A and can get arbitrary large for matrices with large condition numbers. In this case, we speak of *ill-conditioned* problems.

Consider the problem (1.1) with $u \in \mathbb{R}^n$ and $f \in \mathbb{R}^n$ being n-dimensional vectors with real entries and $A \in \mathbb{R}^{n\times n}$ being a matrix with real entries. Assume further A to be symmetric and positive definite.

We know from the spectral theory of symmetric matrices that there exist eigenvalues $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n > 0$ and corresponding (orthonormal) eigenvectors $a_j \in \mathbb{R}^n$ for $j \in \{1,\ldots,n\}$ such that A can be written as

$$A = \sum_{j=1}^n \lambda_j a_j a_j^\top. \tag{1.6}$$

It is well known from numerical linear algebra that the condition number $\kappa = \lambda_1/\lambda_n$ is a measure of how stable (1.1) can be solved, which we will illustrate what follows.

We assume that we measure f_δ instead of f, with $\|f - f_\delta\|_2 \leqslant \delta\|A\| = \delta\lambda_1$, where $\|\cdot\|_2$ denotes the Euclidean norm of $\mathbb{R}^n$ and $\|A\|$ the operator norm of A (which equals the largest eigenvalue of A). Then, if we further denote with u_δ the solution of $Au_\delta = f_\delta$, the difference between u_δ and the solution u to (1.1) is

$$u - u_\delta = \sum_{j=1}^n \lambda_j^{-1} a_j a_j^\top (f - f_\delta).$$

Therefore, we can estimate

$$\|u - u_\delta\|_2^2 = \sum_{j=1}^n \lambda_j^{-2} \underbrace{\|a_j\|_2^2}_{=1} |a_j^\top (f - f_\delta)|^2 \leqslant \lambda_n^{-2}\|f - f_\delta\|_2^2,$$

due to the orthonormality of eigenvectors, the Cauchy-Schwarz inequality, and $\lambda_n \leqslant \lambda_j$. Thus, taking square roots on both sides yields the estimate

$$\|u - u_\delta\|_2 \leqslant \lambda_n^{-1}\|f - f_\delta\|_2 \leqslant \kappa\delta.$$

Hence, we observe that in the worst case an error δ in the data y is amplified by the condition number κ of the matrix A. A matrix with large κ is therefore called *ill-conditioned*. We want to demonstrate the effect of this error amplification with a small example.

Example 1.2.1. Let us consider the matrix

$$A = \begin{pmatrix} 1 & 1 \\ 1 & \frac{1001}{1000} \end{pmatrix},$$

which has eigenvalues $\lambda_j = 1 + \frac{1}{2000} \pm \sqrt{1 + \frac{1}{2000^2}}$, condition number $\kappa \approx 4002 \gg 1$, and operator norm $\|A\| \approx 2$. For given data $f = (1,1)^\top$ the solution to $Au = f$ is $u = (1,0)^\top$.

Now let us instead consider perturbed data $f_\delta = (99/100, 101/100)^\top$. The solution u_δ to $Au_\delta = f_\delta$ is then $u_\delta = (-19.01, 20)^\top$.

Let us reflect on the amplification of the measurement error. By our initial assumption we find that $\delta = \|f - f_\delta\|/\|A\| \approx \|(0.01, -0.01)^\top\|/2 = \sqrt{2}/200$. Moreover, the norm of the error in the reconstruction is then $\|u - u_\delta\| = \|(20.01, 20)^\top\| \approx 20\sqrt{2}$. As a result, the amplification due to the perturbation is $\|u - u_\delta\|/\delta \approx 4000 \approx \kappa$.

1.2.5 Tomography

In almost any tomography application the underlying inverse problem is either the inversion of the Radon transform[1] or of the X-ray transform.

For $u \in C_0^\infty(\mathbb{R}^n)$, $s \in \mathbb{R}$, and $\theta \in S^{n-1}$ the *Radon transform* $R : C_0^\infty(\mathbb{R}^n) \to C^\infty(S^{n-1} \times \mathbb{R})$ can be defined as the integral operator

$$\begin{aligned} f(\theta, s) = (\mathcal{R}u)(\theta, s) &= \int_{x \cdot \theta = s} u(x)\, dx \\ &= \int_{\theta^\perp} u(s\theta + y)\, dy, \end{aligned} \tag{1.7}$$

which, for $n = 2$, coincides with the X-ray transform,

$$f(\theta, s) = (\mathcal{P}u)(\theta, s) = \int_{\mathbb{R}} u(s\theta + t\theta^\perp)\, dt,$$

for $\theta \in S^{n-1}$ and $\theta^\perp$ being the vector orthogonal to θ. Hence, the X-ray transform (and therefore also the Radon transform in two dimensions) integrates the function u over lines in $\mathbb{R}^n$, see Fig. 1.1[2].

Example 1.2.2. Let $n = 2$. Then S^{n-1} is simply the unit sphere $S^1 = \{\theta \in \mathbb{R}^2 \mid \|\theta\| = 1\}$. We can choose for instance $\theta = (\cos(\varphi), \sin(\varphi))^\top$, for $\varphi \in [0, 2\pi)$, and parametrise the Radon transform in terms of φ and s, i.e.

$$f(\varphi, s) = (\mathcal{R}u)(\varphi, s) = \int_{\mathbb{R}} u(s\cos(\varphi) - t\sin(\varphi), s\sin(\varphi) + t\cos(\varphi))\, dt. \tag{1.8}$$

[1] Named after the Austrian mathematician Johann Karl August Radon (16 December 1887 – 25 May 1956).

[2] Figure adapted from Wikipedia `https://commons.wikimedia.org/w/index.php?curid=3001440`, by Begemotv2718, CC BY-SA 3.0.

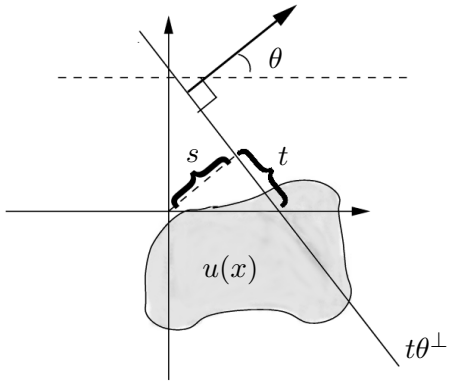


Figure 1.1: Visualization of the Radon transform in two dimensions (which coincides with the X-ray transform). The function u is integrated over the ray parametrized by θ and s.[3]

Note that—with respect to the origin of the reference coordinate system—φ determines the angle of the line along one wants to integrate, while s is the offset from that line from the centre of the coordinate system.

It can be shown that the Radon transform is linear and continuous, i.e. $R \in \mathcal{L}(L^2(B), L^2(Z))$, and even compact.

In **X-ray Computed Tomography (CT)**, the unknown quantity u represents a spatially varying density that is exposed to X-radiation from different angles, and that absorbs the radiation according to its material or biological properties.

The basic modelling assumption for the intensity decay of an X-ray beam is that within a small distance Δt it is proportional to the intensity itself, the density, and the distance, i.e.

$$\frac{I(x + (t + \Delta t)\theta) - I(x + t\theta)}{\Delta t} = -I(x + t\theta)u(x + t\theta),$$

for $x \in \theta^\perp$. By taking the limit $\Delta t \to 0$ we end up with the ordinary differential equation

$$\frac{d}{dt} I(x + t\theta) = -I(x + t\theta)u(x + t\theta), \tag{1.9}$$

Let $R > 0$ be the radius of the domain of interest centred at the origin. Then, we integrate (1.9) from $t = -\sqrt{R^2 - \|x\|_2^2}$, the position of the emitter, to $t = \sqrt{R^2 - \|x\|_2^2}$, the position of the detector, and obtain

$$\int_{-\sqrt{R^2-\|x\|_2^2}}^{\sqrt{R^2-\|x\|_2^2}} \frac{\frac{d}{dt} I(x + t\theta)}{I(x + t\theta)}\, dt = -\int_{-\sqrt{R^2-\|x\|_2^2}}^{\sqrt{R^2-\|x\|_2^2}} u(x + t\theta)\, dt\,.$$

Note that, due to $d/dx \log(f(x)) = f'(x)/f(x)$, the left hand side in the above equation simplifies to

$$\int_{-\sqrt{R^2-\|x\|_2^2}}^{\sqrt{R^2-\|x\|_2^2}} \frac{\frac{d}{dt} I(x + t\theta)}{I(x + t\theta)}\, dt = \log\left(I\left(x + \sqrt{R^2 - \|x\|_2^2}\theta\right)\right) - \log\left(I\left(x - \sqrt{R^2 - \|x\|_2^2}\theta\right)\right).$$

As we know the radiation intensity at both the emitter and the detector, we therefore know $f(x, \theta) = \log(I(x - \theta\sqrt{R^2 - \|x\|_2^2})) - \log(I(x + \theta\sqrt{R^2 - \|x\|_2^2}))$ and we can write the

estimation of the unknown density u as the inverse problem of the X-ray transform (1.8) (if we further assume that u can be continuously extended to zero outside of the circle of radius R).

1.2.6 Groundwater flow/hydraulic tomography

One goal in hydraulic tomography is to estimate the permeability of a groundwater reservoir. The permeability describes the conductivity of the groundwater reservoir and is, e.g., used to estimate the travel time of toxic or radioactive particles in the groundwater.

To estimate the permeability, the water pressure in several position within the reservoir is measured. Pressure head and permeability are linked through Darcy's law and the (assumed) incompressibility of water.

Let $D \subseteq \mathbb{R}^d$ $(d = 1, 2, 3)$ be an open, bounded, connected set with smooth boundary representing the groundwater reservoir. Let $a : \overline{D} \to (0, \infty)$ be a continuously differentiable function representing the permeability and let $s : \overline{D} \to \mathbb{R}$ be a continuous function representing the water sources in the reservoir. Furthermore, assume that the water pressure is 0 outside of D.

Darcy's law states that the pressure $p : D \to \mathbb{R}$, the flux $\vec{q} : D \to \mathbb{R}^d$, and the permeability in the reservoir are related as follows:

$$\vec{q}(x) = -a(x)\nabla p(x) \qquad (x \in D).$$

Incompressibility on the other hand requires that the divergence of the flux is fully controlled by in- and outflow given through the source term s:

$$\nabla \cdot \vec{q}(x) = s(x) \qquad (x \in D).$$

Finally, we can combine these assertions and obtain the elliptic partial differential equation

$$\begin{aligned} -\nabla \cdot a(x)\nabla p(x) &= s(x) && (x \in D) \\ p(x) &= 0 && (x \in \partial D). \end{aligned}$$

In the described set-up, we now observe the pressure p in several positions $x_1, \ldots, x_I \in D$, e.g., we observe $f_n = (p(x_i) : i = 1, ..., I) + n$. We consider the inverse problem consisting in the estimation of the permeability a using the pressure measurements f_n. Indeed, using noisy point evaluations of the solution of the partial differential equation, we try to estimate its diffusion coefficient. Note that the map $a \mapsto (p(x_i) : i = 1, ..., I)$ is non-linear. Hence, this inverse problem is a non-linear inverse problem.

Chapter 2

Generalised Solutions

Functional analysis is the basis of the theory that we will cover in this course. We cannot recall all basic concepts of functional analysis and instead refer to popular textbooks that deal with this subject, e.g., [12, 37, 33]. Nevertheless, we shall recall a few important definitions that will be used in this lecture.

We will focus on inverse problems with *bounded linear operators* A, i.e. $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ with

$$\|A\|_{\mathcal{L}(\mathcal{X},\mathcal{Y})} := \sup_{u \in \mathcal{X}\setminus\{0\}} \frac{\|Au\|_{\mathcal{Y}}}{\|u\|_{\mathcal{X}}} = \sup_{\|u\|_{\mathcal{X}} \leqslant 1} \|Au\|_{\mathcal{Y}} < \infty.$$

For $A\colon \mathcal{X} \to \mathcal{Y}$ we further want to denote by

1. $\mathcal{D}(A) := \mathcal{X}$ the domain,
2. $\mathcal{N}(A) := \{u \in \mathcal{X} \mid Au = 0\}$ the kernel,
3. $\mathcal{R}(A) := \{f \in \mathcal{Y} \mid f = Au, u \in \mathcal{X}\}$ the range

of A.

We say that A is continuous at $u \in \mathcal{X}$ if for all $\varepsilon > 0$ there exists $\delta > 0$ with

$$\|Au - Av\|_{\mathcal{Y}} \leqslant \varepsilon \text{ for all } v \in \mathcal{X} \text{ with } \|u - v\|_{\mathcal{X}} \leqslant \delta.$$

For linear K it can be shown that continuity is equivalent to boundedness, i.e. the existence of a constant $C > 0$ such that

$$\|Au\|_{\mathcal{Y}} \leqslant C\|u\|_{\mathcal{X}}$$

for all $u \in \mathcal{X}$. Note that this constant C actually equals the operator norm $\|A\|_{\mathcal{L}(\mathcal{X},\mathcal{Y})}$.

In this Chapter we only consider $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ with $\mathcal{X}$ and $\mathcal{Y}$ being Hilbert spaces. From functional calculus we know that every Hilbert space $\mathcal{U}$ is equipped with a *scalar product*, which we are going to denote by $\langle \cdot, \cdot \rangle_{\mathcal{U}}$ (or simply $\langle \cdot, \cdot \rangle$, whenever the space is clear from the context). In analogy to the transpose of a matrix, this scalar product structure together with the theorem of Fréchet-Riesz [37, Section 2.10, Theorem 2.E] allows us to define the (unique) *adjoint operator* of A, denoted with A^*, as follows:

$$\langle Au, v \rangle_{\mathcal{Y}} = \langle u, A^* v \rangle_{\mathcal{X}}, \text{ for all } u \in \mathcal{X}, v \in \mathcal{Y}.$$

In addition to that, a scalar product can be used to define orthogonality. Two elements $u, v \in \mathcal{X}$ are said to be *orthogonal* if $\langle u, v\rangle = 0$. For a subset $\mathcal{X}' \subset \mathcal{X}$ the *orthogonal complement* of $\mathcal{X}'$ in $\mathcal{X}$ is defined as

$$\mathcal{X}'^{\perp} \coloneqq \left\{u \in \mathcal{X} \mid \langle u, v\rangle_{\mathcal{X}} = 0 \text{ for all } v \in \mathcal{X}'\right\}.$$

One can show that $\mathcal{X}'^{\perp}$ is a closed subspace and that $\mathcal{X}^{\perp} = \{0\}$. Moreover, we have that $\mathcal{X}' \subset (\mathcal{X}'^{\perp})^{\perp}$. If $\mathcal{X}'$ is a closed subspace then we even have $\mathcal{X}' = (\mathcal{X}'^{\perp})^{\perp}$. In this case there exists the *orthogonal decomposition*

$$\mathcal{X} = \mathcal{X}' \oplus \mathcal{X}'^{\perp},$$

which means that every element $u \in \mathcal{X}$ can uniquely be represented as

$$u = x + x^{\perp} \text{ with } x \in \mathcal{X}' \text{ and } x^{\perp} \in \mathcal{X}'^{\perp},$$

see for instance [37, Section 2.9, Corollary 1].

The mapping $u \mapsto x$ defines a linear operator $P_{\mathcal{X}'} \in \mathcal{L}(\mathcal{X}, \mathcal{X})$ that is called *orthogonal projection* on $\mathcal{X}'$.

Lemma 2.0.1 (cf. [28, Section 5.16])**.** *Let $\mathcal{X}' \subset \mathcal{X}$ be a closed subspace. The orthogonal projection onto $\mathcal{X}'$ satisfies the following conditions:*

1. *$P_{\mathcal{X}'}$ is self-adjoint, i.e. $P^*_{\mathcal{X}'} = P_{\mathcal{X}'}$,*
2. *$\|P_{\mathcal{X}'}\|_{\mathcal{L}(\mathcal{X},\mathcal{X})} = 1$ (if $\mathcal{X}' \neq \{0\}$),*
3. *$I - P_{\mathcal{X}'} = P_{\mathcal{X}'^{\perp}}$,*
4. *$\|u - P_{\mathcal{X}'}u\|_{\mathcal{X}} \leqslant \|u - v\|_{\mathcal{X}}$ for all $v \in \mathcal{X}'$,*
5. *$x = P_{\mathcal{X}'}u$ if and only if $x \in \mathcal{X}'$ and $u - x \in \mathcal{X}'^{\perp}$.*

Remark 2.0.2. Note that for a non-closed subspace $\mathcal{X}'$ we only have $(\mathcal{X}'^{\perp})^{\perp} = \overline{\mathcal{X}'}$. For $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ we therefore have

- $\mathcal{R}(A)^{\perp} = \mathcal{N}(A^*)$ and thus $\mathcal{N}(A^*)^{\perp} = \overline{\mathcal{R}(A)}$,
- $\mathcal{R}(A^*)^{\perp} = \mathcal{N}(A)$ and thus $\mathcal{N}(A)^{\perp} = \overline{\mathcal{R}(A^*)}$.

Hence, we can deduce the following orthogonal decompositions

$$\mathcal{X} = \mathcal{N}(A) \oplus \overline{\mathcal{R}(A^*)} \text{ and } \mathcal{Y} = \mathcal{N}(A^*) \oplus \overline{\mathcal{R}(A)}.$$

We will also need the follwoing relationship between the ranges of A^* and A^*A.

Lemma 2.0.3. *Let $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$. Then $\overline{\mathcal{R}(A^*A)} = \overline{\mathcal{R}(A^*)}$.*

Proof. It is clear that $\overline{\mathcal{R}(A^*A)} = \overline{\mathcal{R}(A^*|_{\mathcal{R}(A)})} \subseteq \overline{\mathcal{R}(A^*)}$, so we are left to prove that $\overline{\mathcal{R}(A^*)} \subseteq \overline{\mathcal{R}(A^*A)}$.

Let $u \in \overline{\mathcal{R}(A^*)}$ and let $\varepsilon > 0$. Then, there exists $f \in \mathcal{N}(A^*)^{\perp} = \overline{\mathcal{R}(A)}$ with $\|A^*f - u\|_{\mathcal{X}} < \varepsilon/2$ (recall the orthogonal decomposition in Remark 2.0.2). As $\mathcal{N}(A^*)^{\perp} = \overline{\mathcal{R}(A)}$, there exists $x \in \mathcal{X}$ such that $\|Ax - f\|_{\mathcal{Y}} < \varepsilon/(2\|A\|_{\mathcal{L}(\mathcal{X},\mathcal{Y})})$. Putting these together we have

$$\begin{aligned}\|A^*Ax - u\|_{\mathcal{X}} &\leqslant \|A^*Ax - A^*f\|_{\mathcal{X}} + \|A^*f - u\|_{\mathcal{X}} \\ &\leqslant \underbrace{\|A^*\|_{\mathcal{L}(\mathcal{Y},\mathcal{X})}\|Ax - f\|_{\mathcal{Y}}}_{<\varepsilon/2} + \underbrace{\|A^*f - u\|_{\mathcal{X}}}_{<\varepsilon/2} < \varepsilon\end{aligned}$$

which shows that $u \in \overline{\mathcal{R}(A^*A)}$ and thus also $\overline{\mathcal{R}(A^*)} \subseteq \overline{\mathcal{R}(A^*A)}$. □

2.1 Generalised Inverses

Recall the inverse problem

$$Au = f, \tag{2.1}$$

where $A\colon \mathcal{X} \to \mathcal{Y}$ is a linear bounded operator and $\mathcal{X}$ and $\mathcal{Y}$ are Hilbert spaces.

Definition 2.1.1 (Minimal-norm solutions)**.** *An element $u \in \mathcal{X}$ is called*

- *a least-squares solution of* (2.1) *if*

$$\|Au - f\|_{\mathcal{Y}} = \inf\{\|Av - f\|_{\mathcal{Y}}, \quad v \in \mathcal{X}\};$$

- *a minimal-norm solution of* (2.1) *(and is denoted by $u^\dagger$) if*

$$\|u^\dagger\|_{\mathcal{X}} \leqslant \|v\|_{\mathcal{X}} \quad \textit{for all least squares solutions } v.$$

Remark 2.1.2. Since $\mathcal{R}(A)$ is not closed in general (it is never closed for a compact operator, unless the range is finite-dimensional), a least-squares solution may not exist. If it exists, then the minimal-norm solution is unique (it is the orthogonal projection of the zero element onto an affine subspace defined by $\|Au - f\|_{\mathcal{Y}} = \min\{\|Av - f\|_{\mathcal{Y}}, \quad v \in \mathcal{X}\}$).

In numerical linear algebra it is a well known fact that the normal equations can be used to compute least-squares solutions. The same holds true in the infinite-dimensional case.

Theorem 2.1.3. *Let $f \in \mathcal{Y}$ and $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$. Then, the following three assertions are equivalent.*

1. *$u \in \mathcal{X}$ satisfies $Au = P_{\overline{\mathcal{R}(A)}} f$.*
2. *u is a least squares solution of the inverse problem* (2.1)*.*
3. *u solves the* normal equation

$$A^*Au = A^*f. \tag{2.2}$$

Remark 2.1.4. The name normal equation is derived from the fact that for any solution u its residual $Au - f$ is orthogonal (normal) to $\mathcal{R}(A)$. This can be readily seen, as we have for any $v \in \mathcal{X}$ that

$$0 = \langle v, A^*(Au - f)\rangle_{\mathcal{X}} = \langle Av, Au - f\rangle_{\mathcal{Y}}$$

which shows $Au - f \in \mathcal{R}(A)^\perp$.

Proof of Theorem 2.1.3. For $1 \Rightarrow 2$: Let $u \in \mathcal{X}$ such that $Au = P_{\overline{\mathcal{R}(A)}} f$ and let $v \in \mathcal{X}$ be arbitrary. With the basic properties of the orthogonal projection, Lemma 2.0.1 4, we have

$$\|Au - f\|_{\mathcal{Y}} = \|ff - P_{\overline{\mathcal{R}(A)}} f\|_{\mathcal{Y}} \leqslant \inf_{g \in \overline{\mathcal{R}(A)}} \|g - f\|_{\mathcal{Y}} \leqslant \inf_{g \in \mathcal{R}(A)} \|g - f\|_{\mathcal{Y}} = \inf_{v \in \mathcal{X}} \|Av - f\|_{\mathcal{Y}},$$

which shows that u is a least squares solution.

For $2 \Rightarrow 3$: Let $u \in \mathcal{X}$ be a least squares solution and let $v \in \mathcal{X}$ an arbitrary element. We define the quadratic polynomial $F\colon \mathbb{R} \to \mathbb{R}$,

$$F(\lambda) := \|A(u + \lambda v) - f\|_{\mathcal{Y}}^2 = \lambda^2 \|Av\|_{\mathcal{Y}}^2 - 2\lambda \langle Av, f - Au\rangle_{\mathcal{Y}} + \|f - Au\|_{\mathcal{Y}}^2 .$$

A necessary condition for $u \in \mathcal{X}$ to be a least squares solution is $F'(0) = 0$, which leads to $\langle v, A^*(f - Au)\rangle_{\mathcal{X}} = 0$. As v was arbitrary, it follows that the normal equation (2.2) must hold.

For $3 \Rightarrow 1$: From the normal equation it follows that $A^*(f - Au) = 0$, which is equivalent to $f - Au \in \mathcal{R}(A)^\perp$, see Remark 2.1.4. Since $\mathcal{R}(A)^\perp = \left(\overline{\mathcal{R}(A)}\right)^\perp$ and $Au \in \mathcal{R}(A) \subset \overline{\mathcal{R}(A)}$, the assertion follows from Lemma 2.0.1 5:

$$Au = P_{\overline{\mathcal{R}(A)}} f \Leftrightarrow Au \in \overline{\mathcal{R}(A)} \text{ and } f - Au \in \left(\overline{\mathcal{R}(A)}\right)^\perp.$$

□

Lemma 2.1.5. *Let $f \in \mathcal{Y}$ and let $\mathbb{L}$ be the set of least squares solutions to the inverse problem* (2.1)*. Then, $\mathbb{L}$ is non-empty if and only if $f \in \mathcal{R}(A) \oplus \mathcal{R}(A)^\perp$.*

Proof. Let $u \in \mathbb{L}$. It is easy to see that $f = Au + (f - Au) \in \mathcal{R}(A) \oplus \mathcal{R}(A)^\perp$ as the normal equations are equivalent to $f - Au \in \mathcal{R}(A)^\perp$.

Consider now $f \in \mathcal{R}(A) \oplus \mathcal{R}(A)^\perp$. Then there exists $u \in \mathcal{X}$ and $g \in \mathcal{R}(A)^\perp = \left(\overline{\mathcal{R}(A)}\right)^\perp$ such that $f = Au + g$ and thus $P_{\overline{\mathcal{R}(A)}} f = P_{\overline{\mathcal{R}(A)}} Au + P_{\overline{\mathcal{R}(A)}} g = Au$ and the assertion follows from Theorem 2.1.3 1. □

Remark 2.1.6. If the dimensions of $\mathcal{X}$ and $\mathcal{R}(A)$ are finite, then $\mathcal{R}(A)$ is closed, i.e. $\overline{\mathcal{R}(A)} = \mathcal{R}(A)$. Thus, in a finite dimensional setting, there always exists a least squares solution.

Theorem 2.1.7. *Let $f \in \mathcal{R}(A) \oplus \mathcal{R}(A)^\perp$. Then there exists a unique minimal norm solution $u^\dagger$ to the inverse problem* (2.1) *and all least squares solutions are given by $\{u^\dagger\} + \mathcal{N}(A)$.*

Proof. From Lemma 2.1.5 we know that there exists a least squares solution. As noted in Remark 2.1.2, in this case the minimal-norm solution is unique. Let φ be an arbitrary least-squares solution. Using Theorem 2.1.3 we get

$$A(\varphi - u^\dagger) = A\varphi - Au^\dagger = P_{\overline{\mathcal{R}(A)}} f - P_{\overline{\mathcal{R}(A)}} f = 0, \tag{2.3}$$

which shows that $\varphi - u^\dagger \in \mathcal{N}(A)$, hence the assertion. □

If a least-squares solution exists for a given $f \in \mathcal{Y}$ then the minimal-norm solution can be computed (at least in theory) using the Moore-Penrose generalised inverse.

Definition 2.1.8. *Let $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ and let*

$$\widetilde{A} \coloneqq A|_{\mathcal{N}(A)^\perp} : \mathcal{N}(A)^\perp \to \mathcal{R}(A)$$

denote the restriction of A to $\mathcal{N}(A)^\perp$. The Moore-Penrose inverse *$A^\dagger$ is defined as the unique linear extension of $\widetilde{A}^{-1}$ to*

$$\mathcal{D}(A^\dagger) = \mathcal{R}(A) \oplus \mathcal{R}(A)^\perp$$

with

$$\mathcal{N}(A^\dagger) = \mathcal{R}(A)^\perp.$$

Remark 2.1.9. Due to the restriction to $\mathcal{N}(A)^\perp$ and $\mathcal{R}(A)$ we have that $\widetilde{A}$ is injective and surjective. Hence, $\widetilde{A}^{-1}$ exists and is linear and – as a consequence – $A^\dagger$ is well-defined on $\mathcal{R}(A)$.

Moreover, due to the orthogonal decomposition $\mathcal{D}(A^\dagger) = \mathcal{R}(A) \oplus \mathcal{R}(A)^\perp$, there exist for arbitrary $f \in \mathcal{D}(A^\dagger)$ elements $f_1 \in \mathcal{R}(A)$ and $f_2 \in \mathcal{R}(A)^\perp$ with $f = f_1 + f_2$. Therefore, we have

$$A^\dagger f = A^\dagger f_1 + A^\dagger f_2 = A^\dagger f_1 = \widetilde{A}^{-1} f_1 = \widetilde{A}^{-1} P_{\overline{\mathcal{R}(A)}} f\,, \tag{2.4}$$

where we used that $f_2 \in \mathcal{R}(A)^\perp = \mathcal{N}(A^\dagger)$. Thus, $A^\dagger$ is well-defined on the entire domain $\mathcal{D}(A^\dagger)$.

Remark 2.1.10. As orthogonal complements are always closed we get that

$$\overline{\mathcal{D}(A^\dagger)} = \overline{\mathcal{R}(A)} \oplus \mathcal{R}(A)^\perp = \mathcal{Y},$$

and hence, $\mathcal{D}(A^\dagger)$ is dense in $\mathcal{Y}$. Thus, if $\mathcal{R}(A)$ is closed it follows that $\mathcal{D}(A^\dagger) = \mathcal{Y}$ and on the other hand, $\mathcal{D}(A^\dagger) = \mathcal{Y}$ implies $\mathcal{R}(A)$ is closed. We note that for ill-posed problems $\mathcal{R}(A)$ is usually not closed; for instance, if A is compact then $\mathcal{R}(A)$ is closed if and only if it is finite-dimensional [1, Ex.1 Section 7.1].

If A is bijective we have that $A^\dagger = A^{-1}$. We also highlight that the extension $A^\dagger$ is not necessarily continuous.

Theorem 2.1.11 ([20, Prop. 2.4])**.** *Let $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$. Then $A^\dagger$ is continuous, i.e. $A^\dagger \in \mathcal{L}(\mathcal{D}(A^\dagger), \mathcal{X})$, if and only if $\mathcal{R}(A)$ is closed.*

Example 2.1.12. To illustrate the definition of the Moore-Penrose inverse we consider a simple example in finite dimensions. Let the linear operator $A\colon \mathbb{R}^3 \to \mathbb{R}^2$ be given by

$$Ax = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2x_1 \\ 0 \end{pmatrix}.$$

It is easy to see that $\mathcal{R}(A) = \{f \in \mathbb{R}^2 \mid f_2 = 0\}$ and $\mathcal{N}(A) = \{x \in \mathbb{R}^3 \mid x_1 = 0\}$. Thus, $\mathcal{N}(A)^\perp = \{x \in \mathbb{R}^3 \mid x_2, x_3 = 0\}$. Therefore, $\widetilde{A}\colon \mathcal{N}(A)^\perp \to \mathcal{R}(A)$, given by $x \mapsto (2x_1, 0)^\top$, is bijective and its inverse $\widetilde{A}^{-1}\colon \mathcal{R}(A) \to \mathcal{N}(A)^\perp$ is given by $f \mapsto (f_1/2, 0, 0)^\top$.

To get the Moore-Penrose inverse $A^\dagger$, we need to extend $\widetilde{A}^{-1}$ to $\mathcal{R}(A) \oplus \mathcal{R}(A)^\perp$ in such a way that $A^\dagger f = 0$ for all $f \in \mathcal{R}(A)^\perp = \{f \in \mathbb{R}^2 \mid f_1 = 0\}$. It is easy to see that the Moore-Penrose inverse $A^\dagger\colon \mathbb{R}^2 \to \mathbb{R}^3$ is given by the following expression

$$A^\dagger f = \begin{pmatrix} 1/2 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \begin{pmatrix} f_1/2 \\ 0 \\ 0 \end{pmatrix}.$$

Let us consider data $\widetilde{f} = (8, 1)^\top \notin \mathcal{R}(A)$. Then, $A^\dagger \widetilde{f} = A^\dagger (8, 1)^\top = (4, 0, 0)^\top$.

It can be shown that $A^\dagger$ can be characterised by the Moore-Penrose equations.

Theorem 2.1.13 ([20, Prop. 2.3])**.** *The Moore-Penrose inverse $A^\dagger$ satisfies $\mathcal{R}(A^\dagger) = \mathcal{N}(A)^\perp$ and the Moore-Penrose equations*

1. $A^\dagger A = P_{\mathcal{N}(A)^\perp}$,
2. $AA^\dagger = P_{\overline{\mathcal{R}(A)}}\Big|_{\mathcal{D}(A^\dagger)}$,
3. $AA^\dagger A = A$,
4. $A^\dagger AA^\dagger = A^\dagger$,

where $P_{\mathcal{N}(A)}$ and $P_{\overline{\mathcal{R}(A)}}$ denote the orthogonal projections on $\mathcal{N}(A)$ and $\overline{\mathcal{R}(A)}$, respectively.

Proof. First, by the definition of the Moore-Penrose inverse we have for any $u \in \mathcal{X}$

$$A^\dagger Au = A^\dagger A(P_{\mathcal{N}(A)}u + P_{\mathcal{N}(A)^\perp}u) = A^\dagger AP_{\mathcal{N}(A)^\perp}u = \widetilde{A}^{-1}AP_{\mathcal{N}(A)^\perp}u = P_{\mathcal{N}(A)^\perp}u,$$

which proves 1. Now, for any $f \in \mathcal{D}(A^\dagger)$ we have (see (2.4))

$$AA^\dagger f = A\widetilde{A}^{-1}P_{\overline{\mathcal{R}(A)}}f = P_{\overline{\mathcal{R}(A)}}f,$$

which proves 2. Applying A to 1., we get 3., and applying $A^\dagger$ to 2., we get 4., which completes the proof. □

Corollary 2.1.14. The Moore-Penrose inverse is uniquely characterised by 1.–2., that is, if a linear operator $B\colon \mathcal{R}(A) \oplus \mathcal{R}(A)^\perp \to \mathcal{N}(A)$ satisfies $BA = P_{\mathcal{N}(A)^\perp}$ and $AB = P_{\overline{\mathcal{R}(A)}}$ then $B = A^\dagger$.

Proof. First we show that $B|_{\mathcal{R}(A)} = \widetilde{A}^{-1}$. Indeed, let $f = Au \in \mathcal{R}(A)$, where $u \in \mathcal{N}(A)^\perp$. Then

$$Bf = BAu = P_{\mathcal{N}(A)^\perp}u = u = \widetilde{A}^{-1}f,$$

where the last equality holds since $\widetilde{A}$ is bijective and hence uniquely invertible.

Now we prove that $B|_{\mathcal{R}(A)^\perp} = 0$. Indeed, for any $f \in \mathcal{R}(A)^\perp$ we have

$$ABf = P_{\overline{\mathcal{R}(A)}}f = 0.$$

Therefore, B is an extension of $\widetilde{A}^{-1}$ to $\mathcal{R}(A) \oplus \mathcal{R}(A)^\perp$ with $\mathcal{N}(B) = \mathcal{R}(A)^\perp$. Since such an extension is unique, $B = A^\dagger$. □

Remark 2.1.15. If an operator B satisfies only $ABA = A$ (resp. $BAB = B$), it is called the *inner inverse* (resp. *outer inverse*) of A.

The next theorem shows that minimal-norm solutions can indeed be computed using the Moore-Penrose generalised inverse.

Theorem 2.1.16. *For each $f \in \mathcal{D}(A^\dagger)$, the minimal norm solution $u^\dagger$ to the inverse problem* (2.1) *is given via*

$$u^\dagger = A^\dagger f.$$

Proof. As $f \in \mathcal{D}(A^\dagger)$, we know from Theorem 2.1.7 that the minimal norm solution $u^\dagger$ exists and is unique. With $u^\dagger \in \mathcal{N}(A)^\perp$, Lemma 2.1.13, and Theorem 2.1.3 we conclude that

$$u^\dagger = (I - P_{\mathcal{N}(A)})u^\dagger = A^\dagger Au^\dagger = A^\dagger P_{\overline{\mathcal{R}(A)}}f = A^\dagger AA^\dagger f = A^\dagger f.$$

□

As a consequence of Theorem 2.1.16 and Theorem 2.1.3, we find that the minimum norm solution $u^\dagger$ of $Au = f$ is a minimum norm solution of the normal equation (2.2), i.e.

$$u^\dagger = (A^*A)^\dagger A^* f.$$

Thus, in order to compute $u^\dagger$ we can equivalently consider finding the minimum norm solution of the normal equation.

2.2 Compact Operators

Definition 2.2.1. *Let $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$. Then A is said to be* compact *if for any bounded set $B \subset \mathcal{X}$ the closure of its image $\overline{A(B)}$ is compact in $\mathcal{Y}$. We denote the space of compact operators by $\mathcal{K}(\mathcal{X}, \mathcal{Y})$.*

Remark 2.2.2. We can equivalently define an operator A to be compact if the image of a bounded sequence $\{u_j\}_{j\in\mathbb{N}} \subset \mathcal{X}$ contains a convergent subsequence $\{Au_{j_k}\}_{k\in\mathbb{N}} \subset \mathcal{Y}$.

Compact operators are very common in inverse problems. In fact, almost all (linear) inverse problems involve the inversion of a compact operator. As the following result shows, compactness of the forward operator is a major source if ill-posedness.

Theorem 2.2.3. *Let $A \in \mathcal{K}(\mathcal{X}, \mathcal{Y})$ with an infinite dimensional range. Then, the Moore-Penrose inverse of A is discontinuous.*

Proof. As the range $\mathcal{R}(A)$ is of infinite dimension, we can conclude that $\mathcal{X}$ and $\mathcal{N}(A)^\perp$ are also infinite dimensional. We can therefore find a sequence $\{u_j\}_{j\in\mathbb{N}}$ with $u_j \in \mathcal{N}(A)^\perp$, $\|u_j\|_\mathcal{X} = 1$ and $\langle u_j, u_k\rangle_\mathcal{X} = 0$ for $j \neq k$. Since A is a compact operator the sequence $f_j = Au_j$ has a convergent subsequence, hence, for all $\delta > 0$ we can find j, k such that $\|f_j - f_k\|_\mathcal{Y} < \delta$. However, we also obtain

$$\begin{aligned}\|A^\dagger f_j - A^\dagger f_k\|_\mathcal{X}^2 &= \|A^\dagger Au_j - A^\dagger Au_k\|_\mathcal{X}^2 \\ &= \|u_j - u_k\|_\mathcal{X}^2 = \|u_j\|_\mathcal{X}^2 - 2\langle u_j, u_k\rangle_\mathcal{X} + \|u_k\|_\mathcal{X}^2 = 2,\end{aligned}$$

which shows that $A^\dagger$ is discontinuous. Here, the second identity follows from Lemma 2.1.13 1 and the fact that $u_j, u_k \in \mathcal{N}(A)^\perp$. □

To have a better understanding of when we have $f \in \overline{\mathcal{R}(A)} \setminus \mathcal{R}(A)$ for compact operators A, we want to consider the singular value decomposition of compact operators.

Singular value decomposition of compact operators

Theorem 2.2.4 ([23, p. 225, Theorem 9.16])**.** *Let $\mathcal{X}$ be a Hilbert space and $A \in \mathcal{K}(\mathcal{X}, \mathcal{X})$ be self-adjoint. Then there exists an orthonormal basis $\{x_j\}_{j\in\mathbb{N}} \subset \mathcal{X}$ of $\overline{\mathcal{R}(A)}$ and a sequence of eigenvalues $\{\lambda_j\}_{j\in\mathbb{N}} \subset \mathbb{R}$ with $|\lambda_1| \geqslant |\lambda_2| \geqslant \ldots > 0$ such that for all $u \in \mathcal{X}$ we have*

$$Au = \sum_{j=1}^{\infty} \lambda_j \langle u, x_j\rangle_\mathcal{X} x_j\,.$$

The sequence $\{\lambda_j\}_{j\in\mathbb{N}}$ is either finite or we have $\lambda_j \to 0$.

Remark 2.2.5. The notation in the theorem above only makes sense if the sequence $\{\lambda_j\}_{j\in\mathbb{N}}$ is infinite. For the case that there are only finitely many λ_j the sum has to be interpreted as a finite sum.

Moreover, as the eigenvalues are sorted by absolute value $|\lambda_j|$, we have $\|A\|_{\mathcal{L}(\mathcal{X},\mathcal{X})} = |\lambda_1|$.

If A is not self-adjoint, the decomposition in Theorem 2.2.4 does not hold any more. Instead, we can consider the so-called *singular value decomposition* of a compact linear operator.

Theorem 2.2.6. *Let $A \in \mathcal{K}(\mathcal{X},\mathcal{Y})$. Then there exists*

1. *a not-necessarily infinite null sequence $\{\sigma_j\}_{j\in\mathbb{N}}$ with $\sigma_1 \geqslant \sigma_2 \geqslant \ldots > 0$,*
2. *an orthonormal basis $\{x_j\}_{j\in\mathbb{N}} \subset \mathcal{X}$ of $\mathcal{N}(A)^\perp$,*
3. *an orthonormal basis $\{y_j\}_{j\in\mathbb{N}} \subset \mathcal{Y}$ of $\overline{\mathcal{R}(A)}$ with*

$$Ax_j = \sigma_j y_j, \quad A^* y_j = \sigma_j x_j, \quad \text{for all } j \in \mathbb{N}. \tag{2.5}$$

Moreover, for all $u \in \mathcal{X}$ we have the representation

$$Au = \sum_{j=1}^{\infty} \sigma_j \langle u, x_j\rangle \, y_j. \tag{2.6}$$

The sequence $\{(\sigma_j, x_j, y_j)\}$ is called singular system *or* singular value decomposition (SVD) *of A.*

For the adjoint operator A^* we have the representation

$$A^* f = \sum_{j=1}^{\infty} \sigma_j \langle f, y_j\rangle \, x_j \quad \forall f \in \mathcal{Y}. \tag{2.7}$$

Proof. Consider $B = A^*A$ and $C = AA^*$. Both B and C are compact, self-adjoint and positive semidefinite, so that by Theorem 2.2.4 both admit a spectral representation and, by positive semidefiniteness, their eigenvalues are positive. Therefore, we can write

$$Cf = \sum_{j=1}^{\infty} \sigma_j^2 \langle f, y_j\rangle \, y_j \quad \forall f \in \mathcal{Y},$$

where $\{y_j\}$ is an orthonormal basis of $\overline{\mathcal{R}(AA^*)} = \overline{\mathcal{R}(A)}$ (Lemma 2.0.3), $\sigma_j > 0$ for all j and $\sigma_j \to 0$ as $j \to \infty$.

Now consider the element $A^* y_j \in \mathcal{X}$. Since σ_j^2 is an eigenvalue of C for the eigenvector y_j, we get that

$$\sigma_j^2 A^* y_j = A^*(\sigma_j^2 y_j) = A^* C y_j = A^* A A^* y_j = B A^* y_j$$

and therefore σ_j^2 is also an eigenvalue of B (for the eigenvector $A^* y_j$). Now we will show that the system $\left\{\frac{A^* y_j}{\sigma_j}\right\}_{j\in\mathbb{N}}$ forms an orthonormal basis of $\overline{\mathcal{R}(A^*)} = \mathcal{N}(A)^\perp$. Indeed, we have

$$\left\langle \frac{A^* y_j}{\sigma_j}, \frac{A^* y_k}{\sigma_k} \right\rangle = \frac{1}{\sigma_j \sigma_k} \langle y_j, AA^* y_k\rangle = \frac{1}{\sigma_j \sigma_k} \langle y_j, \sigma_k^2 y_k\rangle = \begin{cases} 1, & \text{if } j = k, \\ 0, & \text{otherwise.} \end{cases}$$

Hence, $\left\{\frac{A^*y_j}{\sigma_j}\right\}_{j\in\mathbb{N}}$ are orthonormal. It is also clear that they are dense in $\overline{\mathcal{R}(A^*)} = \mathcal{N}(A)^\perp$, hence they form a basis. Therefore, we can choose $\{x_j\}_{j\in\mathbb{N}} = \left\{\frac{A^*y_j}{\sigma_j}\right\}_{j\in\mathbb{N}}$, i.e.

$$x_j = \sigma_j^{-1} A^* y_j$$

and we get (by construction) that

$$A^* y_j = \sigma_j x_j.$$

We also observe that

$$Ax_j = \sigma_j^{-1} AA^* y_j = \sigma_j^{-1}\sigma_j^2 y_j = \sigma_j y_j,$$

which proves (2.5).

Extending the basis $\{x_j\}$ of $\overline{\mathcal{R}(A^*)}$ to a basis $\{\widetilde{x}_j\}$ of $\mathcal{X}$, we expand an arbitrary $u \in \mathcal{X}$ as $u = \sum_{j=1}^{\infty} \langle u, \widetilde{x}_j\rangle \widetilde{x}_j$. Applying A and using the fact that $\mathcal{X} = \mathcal{N}(A)\oplus\overline{\mathcal{R}(A^*)}$ (Remark 2.0.2), we obtain the singular value decomposition (2.6) (and also (2.7) in a similar manner)

$$Au = \sum_{j=1}^{\infty} \sigma_j \langle u, x_j\rangle y_j \quad \forall u \in \mathcal{X}, \quad A^* f = \sum_{j=1}^{\infty} \sigma_j \langle f, y_j\rangle x_j \quad \forall f \in \mathcal{Y}.$$

□

We can now derive a representation of the Moore-Penrose inverse in terms of the singular value decomposition.

Theorem 2.2.7. *Let $A \in \mathcal{K}(\mathcal{X},\mathcal{Y})$ with singular system $\{(\sigma_j, x_j, y_j)\}_{j\in\mathbb{N}}$ and $f \in \mathcal{D}(A^\dagger)$. Then the Moore-Penrose inverse of A can be written as*

$$A^\dagger f = \sum_{j=1}^{\infty} \sigma_j^{-1} \langle f, y_j\rangle x_j\,. \tag{2.8}$$

Proof. We know that, since $f \in \mathcal{D}(A^\dagger)$, $u^\dagger = A^\dagger f$ solves the normal equations

$$A^* A u^\dagger = A^* f.$$

From Theorem 2.2.6 we know that

$$A^* A u^\dagger = \sum_{j=1}^{\infty} \sigma_j^2 \left\langle u^\dagger, x_j\right\rangle x_j, \quad A^* f = \sum_{j=1}^{\infty} \sigma_j \langle f, y_j\rangle x_j, \tag{2.9}$$

which implies that

$$\left\langle u^\dagger, x_j\right\rangle = \sigma_j^{-1} \langle f, y_j\rangle$$

Expanding $u^\dagger \in \mathcal{N}(A)^\perp$ in the basis $\{x_j\}$, we get

$$u^\dagger = \sum_{j=1}^{\infty} \left\langle u^\dagger, x_j\right\rangle x_j = \sum_{j=1}^{\infty} \sigma_j^{-1} \langle f, y_j\rangle x_j = A^\dagger f.$$

□

The representation (2.8) makes it clear again that the Moore-Penrose inverse is unbounded if $\mathcal{R}(A)$ is infinite dimensional. Indeed, taking the sequence y_j we note that $\|A^\dagger y_j\| = \sigma_j^{-1} \to \infty$, although $\|y_j\| = 1$.

The unboundedness of the Moore-Penrose inverse is also reflected in the fact that the series in (2.8) may not converge for a given f. The convergence criterion for the series is called the *Picard criterion.*

Definition 2.2.8. *We say that the data f satisfy the Picard criterion, if*

$$\|A^\dagger f\|^2 = \sum_{j=1}^{\infty} \frac{|\langle f, y_j\rangle|^2}{\sigma_j^2} < \infty. \tag{2.10}$$

Remark 2.2.9. The Picard criterion is a condition on the decay of the coefficients $\langle f, y_j\rangle$. As the singular values σ_j decay to zero as $j \to \infty$, the Picard criterion is only met if the coefficients $\langle f, y_j\rangle$ decay sufficiently fast.

In case the singular system is given by the Fourier basis, then the coefficients $\langle f, y_j\rangle$ are just the Fourier coefficients of f. Therefore, the Picard criterion is a condition on the decay of the Fourier coefficients which is equivalent to the smoothness of f.

It turns our that the Picard criterion also can be used to characterise elements in the range of the forward operator.

Theorem 2.2.10. *Let $A \in \mathcal{K}(\mathcal{X}, \mathcal{Y})$ with singular system $\{(\sigma_j, x_j, y_j)\}_{j\in\mathbb{N}}$, and $f \in \overline{\mathcal{R}(A)}$. Then $f \in \mathcal{R}(A)$ if and only if the* Picard criterion

$$\sum_{j=1}^{\infty} \frac{\left|\langle f, y_j\rangle_{\mathcal{Y}}\right|^2}{\sigma_j^2} < \infty \tag{2.11}$$

is met.

Proof. Let $f \in \mathcal{R}(A)$, thus there is a $u \in \mathcal{X}$ such that $Au = f$. It is easy to see that we have

$$\langle f, y_j\rangle_{\mathcal{Y}} = \langle Au, y_j\rangle_{\mathcal{Y}} = \langle u, A^* y_j\rangle_{\mathcal{X}} = \sigma_j \langle u, x_j\rangle_{\mathcal{X}}$$

and therefore

$$\sum_{j=1}^{\infty} \sigma_j^{-2} |\langle f, y_j\rangle_{\mathcal{Y}}|^2 = \sum_{j=1}^{\infty} |\langle u, x_j\rangle_{\mathcal{X}}|^2 \leqslant \|u\|_{\mathcal{X}}^2 < \infty\,.$$

Now let the Picard criterion (2.11) hold and define $u := \sum_{j=1}^{\infty} \sigma_j^{-1} \langle f, y_j\rangle_{\mathcal{Y}} x_j \in \mathcal{X}$. It is well-defined by the Picard criterion (2.11) and we conclude

$$Au = \sum_{j=1}^{\infty} \sigma_j^{-1} \langle f, y_j\rangle_{\mathcal{Y}} Ax_j = \sum_{j=1}^{\infty} \langle f, y_j\rangle_{\mathcal{Y}} y_j = P_{\overline{\mathcal{R}(A)}} f = f\,,$$

which shows $f \in \mathcal{R}(A)$. □

Although all ill-posed problems are not easy to solve, some are worse than others, depending on how fast the singular values decay to zero.

Definition 2.2.11. *We say that an ill-posed inverse problem* (2.1) *is* mildly ill-posed *if the singular values decay at most with polynomial speed, i.e. there exist* $\gamma, C > 0$ *such that* $\sigma_j \geqslant Cj^{-\gamma}$ *for all* j*. We call the ill-posed inverse problem* severely ill-posed *if its singular values decay faster than with polynomial speed, i.e. for all* $\gamma, C > 0$ *one has that* $\sigma_j \leqslant Cj^{-\gamma}$ *for* j *sufficiently large.*

Example 2.2.12. Let us consider the example of differentiation again, as introduced in Section 1.2.3. The forward operator $A\colon L^2([0,1]) \to L^2([0,1])$ in this problem is given by

$$(Au)(t) = \int_0^t u(s)\,ds = \int_0^1 K(s,t)u(s)\,ds\,,$$

with $K\colon [0,1]\times[0,1] \to \mathbb{R}$ defined as

$$K(s,t) := \begin{cases} 1 & s \leqslant t \\ 0 & \text{else} \end{cases}.$$

This is a special case of the integral operators as introduced in Section 1.2.1. Since the kernel K is square integrable, A is compact.

The adjoint operator A^* is given via

$$(A^*f)(s) = \int_0^1 K(t,s)f(t)\,dt = \int_s^1 v(t)\,dt\,. \tag{2.12}$$

Now we want to compute the eigenvalues and eigenvectors of A^*A, i.e. we look for σ^2 and $x \in L^2([0,1])$ with

$$\sigma^2 x(s) = (A^*Ax)(s) = \int_s^1 \int_0^t x(r)\,dr\,dt\,.$$

We immediately observe $x(1) = 0$ and further

$$\sigma^2 x'(s) = \frac{d}{ds}\int_s^1 \int_0^t x(r)\,dr\,dt = -\int_0^s x(r)\,dr\,,$$

from which we conclude $x'(0) = 0$. Taking the derivative another time thus yields the ordinary differential equation

$$\sigma^2 x''(s) + x(s) = 0\,,$$

for which solutions are of the form

$$x(s) = c_1 \sin(\sigma^{-1}s) + c_2 \cos(\sigma^{-1}s)\,,$$

with some constants c_1, c_2. In order to satisfy the boundary conditions $x(1) = c_1\sin(\sigma^{-1}) + c_2\cos(\sigma^{-1}) = 0$ and $x'(0) = c_1 = 0$, we chose $c_1 = 0$ and σ such that $\cos(\sigma^{-1}) = 0$. Hence, we have

$$\sigma_j = \frac{2}{(2j-1)\pi} \text{ for } j \in \mathbb{N}\,,$$

and by choosing $c_2 = \sqrt{2}$ we obtain the following normalised representation of x_j:

$$x_j(s) = \sqrt{2}\cos\left(\left(j-\frac{1}{2}\right)\pi s\right).$$

According to (2.5) we further obtain

$$y_j(s) = \sigma_j^{-1}(Ax_j)(s) = \left(j-\frac{1}{2}\right)\pi\int_0^s \sqrt{2}\cos\left(\left(j-\frac{1}{2}\right)\pi t\right)\,dt = \sqrt{2}\sin\left(\left(j-\frac{1}{2}\right)\pi s\right),$$

and hence, for $f \in L^2([0,1])$ the Picard criterion becomes

$$2\sum_{j=1}^{\infty}\sigma_j^{-2}\left(\int_0^1 f(s)\sin\left(\sigma_j^{-1}s\right)\,ds\right)^2 < \infty.$$

Expanding f in the basis $\{y_j\}$

$$f(t) = 2\sum_{j=1}^{\infty}\left(\int_0^1 f(s)\sin\left(\sigma_j^{-1}s\right)\,ds\right)\sin\left(\sigma_j^{-1}t\right)$$

and formally differentiating the series, we obtain

$$f'(t) = 2\sum_{j=1}^{\infty}\sigma_j^{-1}\left(\int_0^1 f(s)\sin\left(\sigma_j^{-1}s\right)\,ds\right)\cos\left(\sigma_j^{-1}t\right).$$

Therefore, the Picard criterion is nothing but the condition for the legitimacy of such differentiation, i.e. for the differentiability of the Fourier series by differentiating its components, and it holds if f is differentiable and $f' \in L^2([0,1])$.

From the decay of the singular values we see that this inverse problem is mildly ill-posed.

Example 2.2.13 (Heat equation)**.** Consider the problem of recovering the initial condition u of the heat equation from an observation f of the solution at some time $T > 0$ (see Section 1.2.2). We consider the heat equation on $(0,\pi)\times\mathbb{R}_+$, with Dirichlet boundary conditions

$$\begin{cases} v_t - v_{xx} = 0 & \text{on } (0,\pi)\times\mathbb{R}_+,\\ v(0,t) = v(\pi,t) = 0 & \text{on } \mathbb{R}_+,\\ v(x,T) = f(x) & \text{on } (0,\pi),\\ v(x,0) = u(x) & \text{on } (0,\pi). \end{cases}$$

The solution to the forward problem (determine f given u) is given by

$$f = Au := \sum_{j=1}^{\infty} e^{-j^2T}\widehat{u}_j\sin(jx),$$

where $\widehat{u}_j = \langle u, \sin(j\cdot)\rangle$ are Fourier coefficients of u. Hence, singular values of A are given by

$$\sigma_j = e^{-j^2T}, \quad j\in\mathbb{N},$$

and

$$\frac{1}{\sigma_j} = e^{j^2T}.$$

Singular values decay exponentially and the inverse problem is severely (exponentially) ill-posed.

Chapter 3

Classical Regularisation Theory

3.1 What is Regularisation?

We have seen that the Moore-Penrose inverse $A^\dagger$ is unbounded if $\mathcal{R}(A)$ is not closed. Therefore, given noisy data f_δ such that $\|f_\delta - f\| \leqslant \delta$, we cannot expect convergence $A^\dagger f_\delta \to A^\dagger f$ as $\delta \to 0$. To achieve convergence, we replace $A^\dagger$ with a family of well-posed (bounded) operators R_α with $\alpha = \alpha(\delta, f_\delta)$ and require that $R_{\alpha(\delta,f_\delta)}(f_\delta) \to A^\dagger f$ for all $f \in \mathcal{D}(A^\dagger)$ and all $f_\delta \in \mathcal{Y}$ s.t. $\|f - f_\delta\|_\mathcal{Y} \leqslant \delta$ as $\delta \to 0$.

We remind ourselves that $\mathcal{L}(\mathcal{X}, \mathcal{Y})$ denotes the space of all bounded (equivalently, continuous) operators $\mathcal{X} \to \mathcal{Y}$.

Definition 3.1.1. *Let $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ be a bounded operator. A family $\{R_\alpha\}_{\alpha>0}$ of continuous operators is called* regularisation *(or* regularisation operator*) of $A^\dagger$ if*

$$R_\alpha f \to A^\dagger f = u^\dagger$$

for all $f \in \mathcal{D}(A^\dagger)$ as $\alpha \to 0$.

Definition 3.1.2. *If the family $\{R_\alpha\}_{\alpha>0}$ consists of linear operators, then one speaks of* linear regularisation *of $A^\dagger$.*

Hence, a regularisation is a pointwise approximation of the Moore–Penrose inverse with continuous operators. As in the interesting cases the Moore–Penrose inverse may not be continuous we cannot expect that the norm of R_α stays bounded as $\alpha \to 0$. This is confirmed by the following results (in the linear case).

Theorem 3.1.3 (Banach–Steinhaus e.g. [12, p. 78], [38, p. 173])**.** *Let $\mathcal{X}, \mathcal{Y}$ be Hilbert spaces and $\{A_j\}_{j\in\mathbb{N}} \subset \mathcal{L}(\mathcal{X}, \mathcal{Y})$ a family of point-wise bounded operators, i.e. for all $u \in \mathcal{X}$ there exists a constant $C(u) > 0$ s.t. $\sup_{j\in\mathbb{N}} \|A_j u\|_\mathcal{Y} \leqslant C(u)$. Then*

$$\sup_{j\in\mathbb{N}} \|A_j\|_{\mathcal{L}(\mathcal{X},\mathcal{Y})} < \infty\,.$$

Corollary 3.1.4 ([38, p. 174])**.** Let $\mathcal{X}, \mathcal{Y}$ be Hilbert spaces and $\{A_j\}_{j\in\mathbb{N}} \subset \mathcal{L}(\mathcal{X}, \mathcal{Y})$. Then the following two conditions are equivalent:

1. There exists $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ such that

$$Au = \lim_{j\to\infty} A_j u \quad \text{for all } u \in \mathcal{X}.$$

2. There is a dense subset $\mathcal{X}' \subset \mathcal{X}$ such that $\lim_{j\to\infty} A_j u$ exists for all $u \in \mathcal{X}'$ and

$$\sup_{j\in\mathbb{N}} \|A_j\|_{\mathcal{L}(\mathcal{X},\mathcal{Y})} < \infty\,.$$

Theorem 3.1.5. *Let $\mathcal{X}$, $\mathcal{Y}$ be Hilbert spaces, $A \in \mathcal{L}(\mathcal{X},\mathcal{Y})$ and $\{R_\alpha\}_{\alpha>0}$ a linear regularisation as defined in Definition 3.1.2. If $A^\dagger$ is not continuous, $\{R_\alpha\}_{\alpha>0}$ cannot be uniformly bounded. In particular, there exist $f \in \mathcal{Y}$ and a sequence $\alpha_j \to 0$ such that $\|R_{\alpha_j} f\| \to \infty$ as $j \to \infty$.*

Proof. We prove the theorem by contradiction and assume that $\{R_\alpha\}_{\alpha>0}$ is uniformly bounded. Hence, there exists a constant C with $\|R_\alpha\|_{\mathcal{L}(\mathcal{Y},\mathcal{X})} \leqslant C$ for all $\alpha > 0$. Due to Definition 3.1.1, we have $R_{\alpha_j} \to A^\dagger$ on $\mathcal{D}(A^\dagger)$ for any sequence $\alpha_j \to 0$. Since $\mathcal{D}(A^\dagger)$ is dense in $\mathcal{Y}$, by Corollary 3.1.4 we get that R_{α_j} converges on $\overline{\mathcal{D}(A^\dagger)} = \mathcal{Y}$ and therefore $A^\dagger$ can be extended to a bounded operator on $\mathcal{L}(\mathcal{Y},\mathcal{X})$, which is a contradiction to the assumption that $A^\dagger$ is not continuous (on $\mathcal{D}(A^\dagger)$).

To prove the second statement, assume that for all $f \in \mathcal{Y}$ and any sequence $\alpha_j \to 0$ we have

$$\sup_{j\in\mathbb{N}} \|R_{\alpha_j} f\|_{\mathcal{Y}} \leqslant C(f) < \infty.$$

Then by Theorem 3.1.3 we have that

$$\sup_{j\in\mathbb{N}} \|R_{\alpha_j}\|_{\mathcal{L}(\mathcal{Y},\mathcal{X})} \leqslant C < \infty,$$

which contradicts the first part of the proof. □

With the additional assumption that $\|AR_\alpha\|_{\mathcal{L}(\mathcal{X},\mathcal{X})}$ is bounded, we can even show that $R_\alpha f$ diverges for all $f \notin \mathcal{D}(A^\dagger)$.

Theorem 3.1.6. *Let $A \in \mathcal{L}(\mathcal{X},\mathcal{Y})$ and $\{R_\alpha\}_{\alpha>0}$ be a linear regularisation of $A^\dagger$. If*

$$\sup_{\alpha>0} \|AR_\alpha\|_{\mathcal{L}(\mathcal{X},\mathcal{X})} < \infty\,,$$

then $\|R_\alpha f\|_{\mathcal{X}} \to \infty$ for all $f \notin \mathcal{D}(A^\dagger)$.

Proof. Define $u_\alpha := R_\alpha f$ for $f \notin \mathcal{D}(A^\dagger)$. Assume that there exists a sequence $\alpha_k \to 0$ such that $\|u_{\alpha_k}\|_{\mathcal{X}}$ is uniformly bounded. Since bounded sets in a Hilbert space are weakly pre-compact, there exists a weakly convergent subsequence $u_{\alpha_{k_l}}$ with some limit $u \in \mathcal{X}$, cf. [21, Section 2.2, Theorem 2.1]. As continuous linear operators are also weakly continuous, we further have $Au_{\alpha_{k_l}} \rightharpoonup Au$.

On the other hand, for any $g \in \mathcal{D}(A^\dagger)$ we have that $AR_{\alpha_{k_l}} g \to AA^\dagger g = P_{\overline{\mathcal{R}(A)}} g$ as $l \to \infty$. By Corollary 3.1.4 we then conclude that this also holds for any $f \in \mathcal{Y}$, i.e. also for $f \notin \mathcal{D}(A^\dagger)$. Hence, we get that

$$AR_{\alpha_{k_l}} f \to P_{\overline{\mathcal{R}(A)}} f$$

and (see first part of proof)

$$AR_{\alpha_{k_l}} f = Au_{\alpha_{k_l}} \rightharpoonup Au.$$

Therefore, we get that $Au = P_{\overline{\mathcal{R}(A)}} f$. Since $\mathcal{Y} = \overline{\mathcal{R}(A)} \oplus \mathcal{R}(A)^\perp$, we get that $f \in \mathcal{R}(A) \oplus \mathcal{R}(A)^\perp = \mathcal{D}(A^\dagger)$ in contradiction to the assumption $f \notin \mathcal{D}(A^\dagger)$. □

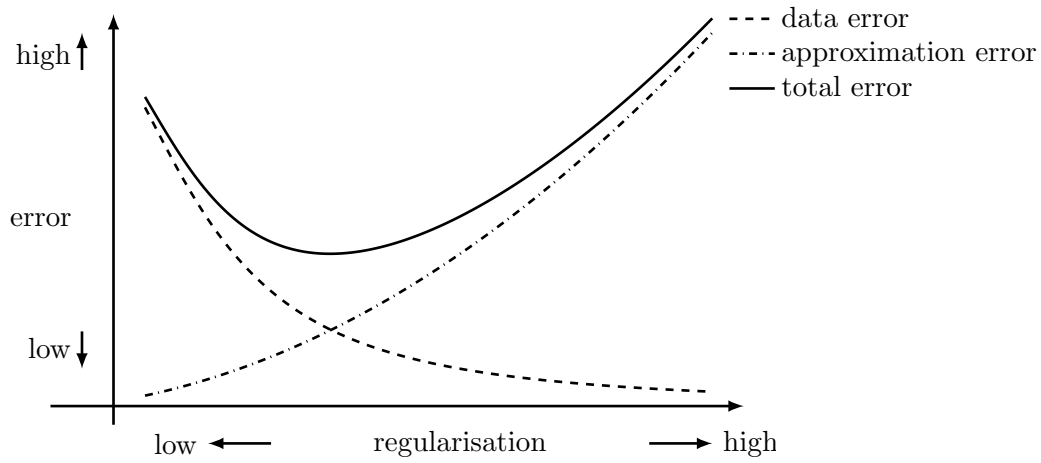


Figure 3.1: The *total error* between a regularised solution and the minimal norm solution decomposes into the *data error* and the *approximation error*. These two errors have opposing trends: For a small regularisation parameter α the error in the data gets amplified through the ill-posedness of the problem and for large α the operator R_α is a poor approximation of the Moore–Penrose inverse.

3.2 Parameter Choice Rules

We have stated in the beginning of this chapter that we would like to obtain a regularisation that would guarantee that $R_\alpha(f_\delta) \to A^\dagger f$ for all $f \in \mathcal{D}(A^\dagger)$ and all $f_\delta \in \mathcal{Y}$ s.t. $\|f - f_\delta\|_\mathcal{Y} \leqslant \delta$ as $\delta \to 0$. This means that the parameter α, referred to as the *regularisation parameter*, needs to be chosen as a function of δ (and perhaps also f_δ) so that $\alpha \to 0$ as $\delta \to 0$ (i.e. we need to regularise less as the data get more precise).

This can be illustrated with the following observation. For linear regularisations we can split the *total error* between the regularised solution of the noisy problem $R_\alpha f_\delta$ and the minimal norm solution of the noise-free problem $u^\dagger = A^\dagger f$ as

$$\begin{aligned} \|R_\alpha f_\delta - u^\dagger\|_\mathcal{X} &\leqslant \|R_\alpha f_\delta - R_\alpha f\|_\mathcal{X} + \|R_\alpha f - u^\dagger\|_\mathcal{X} \\ &\leqslant \underbrace{\delta \|R_\alpha\|_{\mathcal{L}(\mathcal{Y},\mathcal{X})}}_{\text{data error}} + \underbrace{\|R_\alpha f - A^\dagger f\|_\mathcal{X}}_{\text{approximation error}} . \end{aligned} \tag{3.1}$$

The first term of (3.1) is the *data error*; this term unfortunately does not stay bounded for $\alpha \to 0$, which we can conclude from Theorem 3.1.5. The second term, known as the *approximation error*, however vanishes for $\alpha \to 0$, due to the pointwise convergence of R_α to $A^\dagger$. Hence it becomes evident from (3.1) that a good choice of α depends on δ, and needs to be chosen such that the approximation error becomes as small as possible, whilst the data error is being kept at bay. See Figure 3.1 for an illustration.

Parameter choice rules are defined as follows.

Definition 3.2.1. *A function $\alpha\colon \mathbb{R}_{>0} \times \mathcal{Y} \to \mathbb{R}_{>0}$, $(\delta, f_\delta) \mapsto \alpha(\delta, f_\delta)$ is called a* parameter choice rule. *We distinguish between*

1. *a priori parameter choice rules, which depend on δ only;*
2. *a posteriori parameter choice rules, which depend on both δ and f_δ;*

3. heuristic parameter choice rules, which depend on f_δ only.

Now we are ready to define a regularisation that ensures the convergence $R_{\alpha(\delta,f_\delta)}(f_\delta) \to A^\dagger f$ as $\delta \to 0$.

Definition 3.2.2. *Let $\{R_\alpha\}_{\alpha>0}$ be a regularisation of $A^\dagger$. If for all $f \in \mathcal{D}(A^\dagger)$ there exists a parameter choice rule $\alpha : \mathbb{R}_{>0} \times \mathcal{Y} \to \mathbb{R}_{>0}$ such that*

$$\lim_{\delta\to 0} \sup_{f_\delta :\, \|f-f_\delta\|_\mathcal{Y} \leqslant \delta} \|R_\alpha f_\delta - A^\dagger f\|_\mathcal{X} = 0 \tag{3.2}$$

and

$$\lim_{\delta\to 0} \sup_{f_\delta :\, \|f-f_\delta\|_\mathcal{Y} \leqslant \delta} \alpha(\delta, f_\delta) = 0 \tag{3.3}$$

then the pair (R_α, α) is called a convergent regularisation*.*

3.2.1 A priori parameter choice rules

First of all we want to discuss a priori parameter choice rules in more detail. Historically, they were the first to be studied. For every regularisation there exists an a priori parameter choice rule and thus a convergent regularisation.

Theorem 3.2.3 ([20, Prop 3.4])**.** *Let $\{R_\alpha\}_{\alpha>0}$ be a regularisation of $A^\dagger$, for $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$. Then there exists an a priori parameter choice rule $\alpha = \alpha(\delta)$ such that (R_α, α) is a convergent regularisation.*

For linear regularisations, an important characterisation of a priori parameter choice strategies that lead to convergent regularisation methods is as follows.

Theorem 3.2.4. *Let $\{R_\alpha\}_{\alpha>0}$ be a linear regularisation, and $\alpha : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ an a priori parameter choice rule. Then (R_α, α) is a convergent regularisation method if and only if*

a) $\lim_{\delta\to 0} \alpha(\delta) = 0$

b) $\lim_{\delta\to 0} \delta \|R_{\alpha(\delta)}\|_{\mathcal{L}(\mathcal{Y},\mathcal{X})} = 0$

Proof. $\Leftarrow$: Let condition a) and b) be fulfilled. From (3.1) we then observe that for any $f \in \mathcal{D}(A^\dagger)$ and $f_\delta \in \mathcal{Y}$ s.t. $\|f - f_\delta\|_\mathcal{Y} \leqslant \delta$

$$\left\| R_{\alpha(\delta)} f_\delta - A^\dagger f \right\|_\mathcal{X} \to 0 \text{ for } \delta \to 0.$$

Hence, (R_α, α) is a convergent regularisation method.
$\Rightarrow$: Now let (R_α, α) be a convergent regularisation method. We prove that conditions 1 and 2 have to follow from this by showing that violation of either one of them leads to a contradiction to (R_α, α) being a convergent regularisation method. If condition a) is violated, (3.3) is violated and hence, (R_α, α) is not a convergent regularisation method. If condition a) is fulfilled but condition b) is violated, there exists a null sequence $\{\delta_k\}_{k\in\mathbb{N}}$ with $\delta_k \|R_{\alpha(\delta_k)}\|_{\mathcal{L}(\mathcal{Y},\mathcal{X})} \geqslant C > 0$, and hence, we can find a sequence $\{g_k\}_{k\in\mathbb{N}} \subset \mathcal{Y}$ with $\|g_k\|_\mathcal{Y} = 1$ and $\delta_k \|R_{\alpha(\delta_k)} g_k\|_\mathcal{X} \geqslant \widetilde{C}$ for some $\widetilde{C}$. Let $f \in \mathcal{D}(A^\dagger)$ be arbitrary and define $f_k := f + \delta_k g_k$. Then we have on the one hand $\|f - f_k\|_\mathcal{Y} \leqslant \delta_k$, but on the other hand the norm of

$$R_{\alpha(\delta_k)} f_k - A^\dagger f = R_{\alpha(\delta_k)} f - A^\dagger f + \delta_k R_{\alpha(\delta_k)} g_k$$

cannot converge to zero, as the second term $\delta_k R_{\alpha(\delta_k)} g_k$ is bounded from below by a positive constant C by construction. Hence, (3.2) is violated for $f_\delta = f + \delta_k g_k$ and thus, (R_α, α) is not a convergent regularisation method. □

3.2.2 A posteriori parameter choice rules

It is easy to convince oneself that if an a priori parameter choice rule $\alpha = \alpha(\delta)$ defines a convergence regularisation then $\widetilde{\alpha} = \alpha(C\delta)$ with any $C > 0$ also defines a convergent regularisation (for linear regularisations, it is a trivial corollary of Theorem 3.2.4). Therefore, from the asymptotic point of view, all these regularisations are equivalent. For a fixed error level δ, however, they can produce very different solutions. Since in practice we have to deal with a typically small, but fixed δ, we would like to have a parameter choice rule that is sensitive to this value. To achieve this, we need to use more information than merely the error level δ to choose the parameter α and we will obtain this information from the approximate data f_δ.

The basic idea is as follows. Let $f \in \mathcal{D}(A^\dagger)$ and $f_\delta \in \mathcal{Y}$ such that $\|f - f_\delta\| \leqslant \delta$ and consider the *residual* between f_δ and $u_\alpha := R_\alpha f_\delta$, i.e.

$$\|Au_\alpha - f_\delta\|.$$

Let $u^\dagger$ be the minimal norm solution and define

$$\mu := \inf\{\|Au - f\|,\ u \in \mathcal{X}\} = \|Au^\dagger - f\|.$$

We observe that $u^\dagger$ satisfies the following inequality

$$\|Au^\dagger - f_\delta\| \leqslant \|Au^\dagger - f\| + \|f_\delta - f\| \leqslant \mu + \delta$$

and in some cases this estimate may be sharp. Hence, it appears not to be useful to choose $\alpha(\delta, f_\delta)$ with $\|Au_\alpha - f_\delta\| < \mu + \delta$. In general, it may be not straightforward to estimate μ, but if $\mathcal{R}(A)$ is dense in $\mathcal{Y}$, we get that $\mathcal{R}(A)^\perp = \{0\}$ due to Remark 2.0.2 and $\mu = 0$. Therefore, we ideally ensure that $\mathcal{R}(A)$ is dense.

These observations motivate the Morozov's discrepancy principle, which in the case $\mu = 0$ reads as follows.

Definition 3.2.5 (Morozov's discrepancy principle)**.** *Let $u_\alpha = R_\alpha f_\delta$ with $\alpha(\delta, f_\delta)$ chosen as follows*

$$\alpha(\delta, f_\delta) = \sup\{\alpha > 0 \mid \|Au_\alpha - f_\delta\| \leqslant \eta\delta\} \tag{3.4}$$

for given δ, f_δ and a fixed constant $\eta > 1$. Then $u_{\alpha(\delta,f_\delta)} = R_{\alpha(\delta,f_\delta)} f_\delta$ is said to satisfy Morozov's discrepancy principle.

It can be shown that the a-posteriori parameter choice rule (3.4) indeed yields a convergent regularization method [20, Chapter 4.3].

3.2.3 Heuristic parameter choice rules

As the measurement error δ is not always easy to obtain in practice, it is tempting to use a parameter choice rule that only depends on the measured data f_δ and not on their error δ, i.e. to use a heuristic parameter choice rule. Unfortunately, heuristic rules yield convergent regularisations only for well-posed problems, as the following result, known as the Bakushinskii veto [7], demonstrates.

Theorem 3.2.6 ([20, Thm 3.3])**.** *Let $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ and $\{R_\alpha\}$ be a regularization for $A^\dagger$. Let $\alpha = \alpha(f_\delta)$ be a parameter choice rule such that (R_α, α) is a convergent regularization. Then $A^\dagger$ is continuous from $\mathcal{Y}$ to $\mathcal{X}$.*

3.3 Spectral Regularisation

Recall the spectral representation (2.8) of the Moore-Penrose inverse $A^\dagger$

$$A^\dagger f = \sum_{j=1}^{\infty} \frac{1}{\sigma_j} \langle f, y_j \rangle \, x_j \,,$$

where $\{(\sigma_j, x_j, y_j)\}$ is the singular system of A.

The source of ill-posedness of $A^\dagger$ are the eigenvalues $1/\sigma_j$, which explode as $j \to \infty$, since $\sigma_j \to 0$ as $j \to \infty$. Let us construct a regularisation by modifying these eigenvalues as follows

$$R_\alpha f := \sum_{j=1}^{\infty} g_\alpha(\sigma_j) \langle f, y_j \rangle \, x_j \,, \;\; f \in \mathcal{Y}, \tag{3.5}$$

with an appropriate function $g_\alpha \colon \mathbb{R}_+ \to \mathbb{R}_+$ such that $g_\alpha(\sigma) \to \frac{1}{\sigma}$ as $\alpha \to 0$ for all $\sigma > 0$ and

$$g_\alpha(\sigma) \leqslant C_\alpha \text{ for all } \sigma \in \mathbb{R}_+. \tag{3.6}$$

Theorem 3.3.1. *Let $g_\alpha : \mathbb{R}_+ \to \mathbb{R}_+$ be a piecewise continuous function satisfying* (3.6)*, $\lim_{\alpha \to 0} g_\alpha(\sigma) = \frac{1}{\sigma}$ and*

$$\sup_{\alpha, \sigma} \sigma g_\alpha(\sigma) \leqslant \gamma \tag{3.7}$$

for some constant $\gamma > 0$. If R_α is defined as in (3.5)*, we have*

$$R_\alpha f \to A^\dagger f \;\; \text{as } \alpha \to 0$$

for all $f \in \mathcal{D}(A^\dagger)$.

Proof. From the singular value decomposition of $A^\dagger$ and the definition of R_α we obtain

$$R_\alpha f - A^\dagger f = \sum_{j=1}^{\infty} \left(g_\alpha(\sigma_j) - \frac{1}{\sigma_j} \right) \langle f, y_j \rangle_{\mathcal{Y}} \, x_j = \sum_{j=1}^{\infty} (\sigma_j g_\alpha(\sigma_j) - 1) \, \langle u^\dagger, x_j \rangle_{\mathcal{X}} \, x_j \,.$$

Consider

$$\| R_\alpha f - A^\dagger f \|_{\mathcal{X}}^2 = \sum_{j=1}^{\infty} (\sigma_j g_\alpha(\sigma_j) - 1)^2 \left| \langle u^\dagger, x_j \rangle_{\mathcal{X}} \right|^2 .$$

From (3.7) we can conclude

$$(\sigma_j g_\alpha(\sigma_j) - 1)^2 \leqslant (1 + \gamma^2) \,,$$

whilst

$$\sum_{j=1}^{\infty} (1 + \gamma^2) \left| \langle u^\dagger, x_j \rangle_{\mathcal{X}} \right|^2 = (1 + \gamma^2) \| u^\dagger \|^2 < +\infty.$$

Therefore, by the reverse Fatou lemma we get the following estimate

$$\begin{aligned}\limsup_{\alpha\to 0}\left\|R_\alpha f - A^\dagger f\right\|_{\mathcal{X}}^2 &= \limsup_{\alpha\to 0}\sum_{j=1}^{\infty}\left(\sigma_j g_\alpha(\sigma_j)-1\right)^2\left(\langle u^\dagger, x_j\rangle_{\mathcal{X}}\right)^2 \\ &\leqslant \sum_{j=1}^{\infty}\left(\limsup_{\alpha\to 0}\sigma_j g_\alpha(\sigma_j)-1\right)^2\left|\langle u^\dagger, x_j\rangle_{\mathcal{X}}\right|^2 = 0\,,\end{aligned}$$

where the last equality is due to the pointwise convergence of $g_\alpha(\sigma_j)$ to $1/\sigma_j$. Hence, we have $\left\|R_\alpha f - A^\dagger f\right\|_{\mathcal{X}} \to 0$ for $\alpha \to 0$ for all $f \in \mathcal{D}(A^\dagger)$. □

Theorem 3.3.2. *Let the assumptions of Theorem 3.3.1 hold and let $\alpha = \alpha(\delta)$ be an a-priori parameter choice rule. Then $(R_{\alpha(\delta)}, \alpha(\delta))$ with R_α as defined in (3.5) is a convergent regularisation method if*

$$\lim_{\delta\to 0}\delta C_{\alpha(\delta)} = 0.$$

Proof. The result follows immediately from $\|R_{\alpha(\delta)}\|_{\mathcal{L}(\mathcal{X},\mathcal{Y})} \leqslant C_{\alpha(\delta)}$ and Theorem 3.2.4. □

3.3.1 Truncated singular value decomposition

As a first example for a spectral regularisation of the form (3.5) we want to consider the so-called *truncated singular value decomposition*. The idea is to discard all singular values below a certain threshold α, which is achieved using the following function g_α

$$g_\alpha(\sigma) = \begin{cases}\frac{1}{\sigma} & \sigma \geqslant \alpha \\ 0 & \sigma < \alpha\end{cases}. \tag{3.8}$$

Note that for all $\sigma > 0$ we naturally obtain $\lim_{\alpha\to 0} g_\alpha(\sigma) = 1/\sigma$. Condition (3.7) is obviously satisfied with $\gamma = 1$ and condition (3.6) with $C_\alpha = \frac{1}{\alpha}$. Therefore, truncated SVD is a convergent regularisation if

$$\lim_{\delta\to 0}\frac{\delta}{\alpha} = 0. \tag{3.9}$$

Equation (3.5) then reads as follows

$$R_\alpha f = \sum_{\sigma_j \geqslant \alpha}\frac{1}{\sigma_j}\langle f, y_j\rangle_{\mathcal{Y}}\, x_j\,, \tag{3.10}$$

for all $f \in \mathcal{Y}$. Note that the sum in (3.10) is always well-defined (i.e. finite) for any $\alpha > 0$ as zero is the only accumulation point of singular vectors of compact operators.

Let $A \in \mathcal{K}(\mathcal{X},\mathcal{Y})$ with singular system $\{(\sigma_j, x_j, y_j)\}_{j\in\mathbb{N}}$, and choose for $\delta > 0$ an index function $j^* : \mathbb{R}_+ \to \mathbb{N}$ with $j^*(\delta) \to \infty$ for $\delta \to 0$ and $\lim_{\delta\to 0}\delta/\sigma_{j^*(\delta)} = 0$. We can then choose $\alpha(\delta) = \sigma_{j^*(\delta)}$ as an a-priori parameter choice rule to obtain a convergent regularisation.

Note that in practice a larger δ implies that more and more singular values have to be cut off in order to guarantee a stable recovery that successfully suppresses the data error.

A disadvantage of this approach is that it requires the knowledge of the singular vectors of A (only finitely many, but the number can still be large).

3.3.2 Tikhonov regularisation

The main idea behind Tikhonov regularisation[1] is to consider the normal equations and shift the eigenvalues of A^*A by a constant factor, which will be associated with the regularisation parameter α. This shift can be realised via the function

$$g_\alpha(\sigma) = \frac{\sigma}{\sigma^2 + \alpha} \tag{3.11}$$

and the corresponding Tikhonov regularisation (3.5) reads as follows

$$R_\alpha f = \sum_{j=1}^{\infty} \frac{\sigma_j}{\sigma_j^2 + \alpha} \langle f, y_j \rangle_{\mathcal{Y}} \, x_j \,. \tag{3.12}$$

Again, we immediately observe that for all $\sigma > 0$ we have $\lim_{\alpha \to 0} g_\alpha(\sigma) = 1/\sigma$. Condition (3.7) is satisfied with $\gamma = 1$. Since $0 \leqslant (\sigma - \sqrt{\alpha})^2 = \sigma^2 - 2\sigma\sqrt{\alpha} + \alpha$, we get that $\sigma^2 + \alpha \geqslant 2\sigma\sqrt{\alpha}$ and

$$\frac{\sigma}{\sigma^2 + \alpha} \leqslant \frac{1}{2\sqrt{\alpha}}.$$

This estimate implies that (3.6) holds with $C_\alpha = \frac{1}{2\sqrt{\alpha}}$. Therefore, Tikhonov regularisation is a convergent regularisation if

$$\lim_{\delta \to 0} \frac{\delta}{\sqrt{\alpha}} = 0. \tag{3.13}$$

The formula (3.12) suggests that we need all singular vectors of A in order to compute the regularisation. However, we note that σ_j^2 are the eigenvalues of A^*A and, hence, $\sigma_j^2 + \alpha$ are the eigenvectors of $A^*A + \alpha I$ (where I is the identity operator). Applying this operator to the regularised solution $u_\alpha = R_\alpha f$, we get

$$(A^*A + \alpha I)u_\alpha = \sum_{j=1}^{\infty} (\sigma_j^2 + \alpha) \langle u_\alpha, x_j \rangle_{\mathcal{X}} \, x_j = \sum_{j=1}^{\infty} (\sigma_j^2 + \alpha) \frac{\sigma_j}{\sigma_j^2 + \alpha} \langle f, y_j \rangle_{\mathcal{Y}} \, x_j = A^* f.$$

Therefore, the regularised solution u_α can be computed without knowing the singular system of A by solving the following well-posed linear equation

$$(A^*A + \alpha I)u_\alpha = A^* f. \tag{3.14}$$

Remark 3.3.3. Rewriting equation (3.14) as

$$A^*(Au_\alpha - f) + \alpha u_\alpha = 0,$$

we note that it looks like a condition for the minimum of some quadratic form. Indeed, it can be easily checked that (3.14) is the first order optimality condition for the following optimisation problem

$$\min_{u \in \mathcal{X}} \frac{1}{2} \|Au - f\|^2 + \alpha \|u\|^2. \tag{3.15}$$

The condition (3.14) is necessary (and, by convexity, sufficient) for the minimum of the functional in (3.15). Therefore, the regularised solution u_α can also be computed by solving (numerically) the variational problem (3.15). This is the starting point for modern variational regularisation methods, which we will consider in the next chapter.

[1] Named after the Russian mathematician Andrey Nikolayevich Tikhonov (30 October 1906 - 7 October 1993)

Chapter 4

Variational Regularisation

Recall the variation formulation of Tikhonov regularisation for some data $f_\delta \in \mathcal{Y}$

$$\min_{u\in\mathcal{X}} \|Au - f_\delta\|^2 + \alpha\|u\|^2.$$

The first term in this expression, $\|Au - f_\delta\|^2$, penalises the misfit between the predictions of the operator A and the measured data f_δ and is called the *fidelity function* or *fidelity term*. The second term, $\|u\|^2$ penalises some unwanted features of the solution (in this case, a large norm) and is called the *regularistaion term*. The regularisation parameter α in this context balances the influence of these two terms on the functional to be minimised.

More generally, using the notation $\mathcal{J}(u)$ for the regulariser, we can formally write down the variational regularisation problem as follows

$$\min_{u\in\mathcal{X}} \frac{1}{2}\|Au - f_\delta\|^2 + \alpha\mathcal{J}(u), \tag{4.1}$$

(the $\frac{1}{2}$ in front of the fidelity term is there to simplify notation later). The regularisation operator R_α is defined as follows

$$R_\alpha f_\delta \in \arg\min_{u\in\mathcal{X}} \frac{1}{2}\|Au - f_\delta\|^2 + \alpha\mathcal{J}(u).$$

In general, the minimiser doesn't have to unique, hence the inclusion and not equality. Other fidelity terms (not just $\|Au - f_\delta\|^2$) are possible and useful in many situations. In this course, however, we will use the squared norm for the sake of simplicity.

In this chapter, we will study the properties of (4.1) for different choices of $\mathcal{J}$, but before that we will recall some necessary theoretical concepts.

4.1 Background

4.1.1 Banach spaces and weak convergence

Banach spaces are complete, normed vector spaces (as Hilbert spaces) but they may not have an inner product. For every Banach space $\mathcal{X}$, we can define the space of linear and continuous functionals which is called the *dual space* $\mathcal{X}^*$ of $\mathcal{X}$, i.e. $\mathcal{X}^* := \mathcal{L}(\mathcal{X}, \mathbb{R})$. Let $u \in \mathcal{X}$ and $p \in \mathcal{X}^*$, then we usually write the *dual product* $\langle p, u\rangle$ instead of $p(u)$. Moreover,

for any $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ there exists a unique operator $A^*\colon \mathcal{Y}^* \to \mathcal{X}^*$, called the *adjoint* of A such that for all $u \in \mathcal{X}$ and $p \in \mathcal{Y}^*$ we have

$$\langle A^* p, u \rangle = \langle p, Au \rangle \ .$$

It is easy to see that either side of the equation are well-defined, e.g. $A^*p \in \mathcal{X}^*$ and $u \in \mathcal{X}$.

The dual space of a Banach space $\mathcal{X}$ can be equipped with the following norm

$$\|p\|_{\mathcal{X}^*} = \sup_{u \in \mathcal{X}, \|u\|_{\mathcal{X}} \leqslant 1} \langle p, u \rangle \, .$$

With this norm the dual space is itself a Banach space. Therefore, it has a dual space as well which we will call the bi-dual space of $\mathcal{X}$ and denote it with $\mathcal{X}^{**} := (\mathcal{X}^*)^*$. As every $u \in \mathcal{X}$ defines a continuous and linear mapping on the dual space $\mathcal{X}^*$ by

$$\langle E(u), p \rangle := \langle p, u \rangle \ ,$$

the mapping $E\colon \mathcal{X} \to \mathcal{X}^{**}$ is well-defined. It can be shown that E is a linear and continuous isometry (and thus injective). In the special case when E is surjective, we call $\mathcal{X}$ *reflexive*. Examples of reflexive Banach spaces include Hilbert spaces and L^q, ℓ^q spaces with $1 < q < \infty$. We call the space $\mathcal{X}$ *separable* if there exists a set $\mathcal{X}' \subset \mathcal{X}$ of at most countable cardinality such that $\overline{\mathcal{X}'} = \mathcal{X}$.

A problem in infinite dimensional spaces is that bounded sequences may fail to have convergent subsequences. An example is for instance in ℓ^2 the sequence $\{u^k\}_{k \in \mathbb{N}} \subset \ell^2, u^k_j = 1$ if $k = j$ and 0 otherwise. It is easy to see that $\|u^k\|_{\ell^2} = 1$ and that there is no $u \in \ell^2$ such that $u^k \to u$. To circumvent this problem, we define a weaker topology on $\mathcal{X}$. We say that $\{u^k\}_{k \in \mathbb{N}} \subset \mathcal{X}$ *converges weakly* to $u \in \mathcal{X}$ if and only if for all $p \in \mathcal{X}^*$ the sequence of real numbers $\{\langle p, u^k \rangle\}_{k \in \mathbb{N}}$ converges and

$$\langle p, u_j \rangle \to \langle p, u \rangle \ .$$

We will denote weak convergence by $u^k \rightharpoonup u$. On a dual space $\mathcal{X}^*$ we could define another topology (in addition to the strong topology induced by the norm and the weak topology as the dual space is a Banach space as well). We say a sequence $\{p^k\}_{k \in \mathbb{N}} \subset \mathcal{X}^*$ *converges weakly-*$*$ to $p \in \mathcal{X}^*$ if and only if

$$\left\langle p^k, u \right\rangle \to \langle p, u \rangle \quad \text{for all } u \in \mathcal{X}$$

and we denote weak-$*$ convergence by $p^k \rightharpoonup^* p$. Similarly, for any topology τ on $\mathcal{X}$ we denote the convergence in that topology by $u^k \xrightarrow{\tau} u$.

With these two new notions of convergence, we can solve the problem of bounded sequences:

Theorem 4.1.1 (Banach-Alaoglu Theorem, e.g. [32, p. 70] or [36, p. 141])**.** *Let $\mathcal{X} = (\mathcal{X}^\diamond)^*$ be the dual of a Banach space $\mathcal{X}^\diamond$. Then the unit ball $\mathcal{B}_{\mathcal{X}} = \{u \in \mathcal{X}\colon \|x\| \leqslant 1\}$ is compact in the weak-$*$ topology. If $\mathcal{X}^\diamond$ is separable, then the weak-$*$ topology is metrisable on bounded sets and every bounded sequence $\{u^k\}_{k \in \mathbb{N}} \subset \mathcal{X}$ has a weak-* convergent subsequence.*

Theorem 4.1.2 ([38, p. 64])**.** *Each bounded sequence $\{u^k\}_{k \in \mathbb{N}}$ in a reflexive separable Banach space $\mathcal{X}$ has a weakly convergent subsequence.*

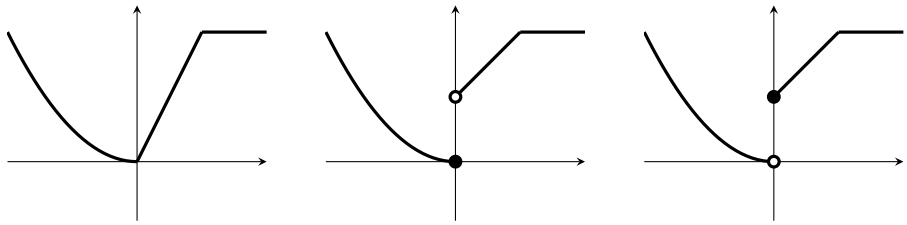


Figure 4.1: Visualisation of lower semi-continuity. The solid dot at a jump indicates the value that the function takes. The function on the left is continuous and thus lower semi-continuous. The functions in the middle and on the right are discontinuous. While the function in the middle is lower semi-continuous, the function on the right is not (due to the limit from the left at the discontinuity).

An important property of functionals, which we will need later, is sequential lower semicontinuity. Roughly speaking this means that the functional values for arguments near an argument u are either close to $E(u)$ or greater than $E(u)$.

Definition 4.1.3. *Let $\mathcal{X}$ be a Banach space with topology $\tau_{\mathcal{X}}$. The functional $E\colon \mathcal{X} \to \bar{\mathbb{R}}$ is said to be* sequentially lower semi-continuous with respect to $\tau_{\mathcal{X}}$ *($\tau_{\mathcal{X}}$-l.s.c.) at $u \in \mathcal{X}$ if*

$$E(u) \leqslant \liminf_{j\to\infty} E(u_j)$$

for all sequences $\{u_j\}_{j\in\mathbb{N}} \subset \mathcal{X}$ with $u_j \to u$ in the topology $\tau_{\mathcal{X}}$ of $\mathcal{X}$.

Remark 4.1.4. For topologies that are not induced by a metric we have to differ between a topological property and its sequential version, e.g. continuous and sequentially continuous. If the topology is induced by a metric, then these two are the same. However, for instance the weak and weak-$*$ topology are generally not induced by a metric (but this is true on bounded sets).

Example 4.1.5. The functional $\|\cdot\|_1 : \ell^2 \to \bar{\mathbb{R}}$ with

$$\|u\|_1 = \begin{cases} \sum_{j=1}^{\infty} |u_j| & \text{if } u \in \ell^1 \\ \infty & \text{else} \end{cases}$$

is weakly (and, hence, strongly) lower semi-continuous in ℓ^2.

Proof. Let $\{u^j\}_{j\in\mathbb{N}} \subset \ell^2$ be a weakly convergent sequence with $u^j \rightharpoonup u \in \ell^2$. We have with $\delta_k : \ell^2 \to \mathbb{R}, \langle \delta_k, v\rangle = v_k$ that for all $k \in \mathbb{N}$

$$u_k^j = \langle \delta_k, u^j\rangle \to \langle \delta_k, u\rangle = u_k\,.$$

The assertion follows then with Fatou's lemma

$$\|u\|_1 = \sum_{k=1}^{\infty} |u_k| = \sum_{k=1}^{\infty} \lim_{j\to\infty} |u_k^j| \leqslant \liminf_{j\to\infty} \sum_{k=1}^{\infty} |u_k^j| = \liminf_{j\to\infty} \|u^j\|_1\,.$$

Note that it is not clear whether both the left and the right hand side are finite. □

4.1.2 Convex analysis

Infinity calculus

We will look at functionals $E : \mathcal{X} \to \bar{\mathbb{R}}$ whose range is modelled to be the *extended real line* $\bar{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$ where the symbol $+\infty$ denotes an element that is not part of the real line that is by definition larger than any other element of the reals, i.e.

$$x < +\infty$$

for all $x \in \mathbb{R}$ (similarly, $x > -\infty$ for all $x \in \mathbb{R}$). This is useful to model constraints: for instance, if we were trying to minimise $E : [-1, \infty) \to \mathbb{R}, x \mapsto x^2$ we could remodel this minimisation problem by $\widetilde{E} : \mathbb{R} \to \bar{\mathbb{R}}$

$$\widetilde{E}(x) = \begin{cases} x^2 & \text{if } x \geqslant -1 \\ \infty & \text{else} \end{cases}.$$

Obviously both functionals have the same minimiser but $\widetilde{E}$ is defined on a vector space and not only on a subset. This has two important consequences: on the on hand, it makes many theoretical arguments easier as we do not need to worry whether $E(x + y)$ is defined or not. On the other hand, it makes practical implementations easier as we are dealing with unconstrained optimisation instead of constrained optimisation. This comes at a cost that some algorithms are not applicable any more, e.g. the function $\widetilde{E}$ is not differentiable everywhere whereas E is (in the interior of its domain).

It is useful to note that one can calculate on the extended real line $\bar{\mathbb{R}}$ as we are used to on the real line $\mathbb{R}$ but the operations with $\pm\infty$ need yet to be defined.

Definition 4.1.6. *The extended real line is defined as* $\bar{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$ *with the following rules that hold for any* $x \in \mathbb{R}$ *and* $\lambda > 0$*:*

$$\begin{aligned} x \pm \infty &:= \pm\infty + x := \pm\infty \\ \lambda \cdot (\pm\infty) &:= \pm\infty \cdot \lambda := \pm\infty, \quad -1 \cdot (\pm\infty) := \mp\infty \\ x/(\pm\infty) &:= 0 \\ \infty + \infty &:= \infty, \quad -\infty - \infty := -\infty\,. \end{aligned}$$

Some calculations are *not defined*, e.g.,

$$+\infty - \infty \,\text{and} \quad (\pm\infty) \cdot (\pm\infty)\,.$$

Using functions with values on the extended real line, one can easily describe sets $\mathcal{C} \subset \mathcal{X}$.

Definition 4.1.7 (Characteristic function)**.** *Let* $\mathcal{C} \subset \mathcal{X}$ *be a set. The function* $\chi_{\mathcal{C}} \colon \mathcal{X} \to \bar{\mathbb{R}}$*,*

$$\chi_{\mathcal{C}}(u) = \begin{cases} 0 & u \in \mathcal{C} \\ \infty & u \in \mathcal{X} \setminus \mathcal{C} \end{cases}$$

is called the characteristic function of the set $\mathcal{C}$*.*

Using characteristic functions, one can easily write constrained optimisation problems as unconstrained ones:

$$\min_{u \in \mathcal{C}} E(u) \quad \Leftrightarrow \quad \min_{u \in \mathcal{X}} E(u) + \chi_{\mathcal{C}}(u).$$

Definition 4.1.8. *Let $\mathcal{X}$ be a vector space and $E\colon \mathcal{X} \to \bar{\mathbb{R}}$ a functional. Then the* effective domain *of E is*

$$\operatorname{dom}(E) := \{u \in \mathcal{X} \mid E(u) < \infty\} .$$

Definition 4.1.9. *A functional E is called* proper *if the effective domain* $\operatorname{dom}(E)$ *is not empty.*

Convexity

A property of fundamental importance of sets and functions is convexity.

Definition 4.1.10. *Let $\mathcal{X}$ be a vector space. A subset $\mathcal{C} \subset \mathcal{X}$ is called* convex*, if $\lambda u + (1-\lambda)v \in \mathcal{C}$ for all $\lambda \in (0,1)$ and all $u, v \in \mathcal{C}$.*

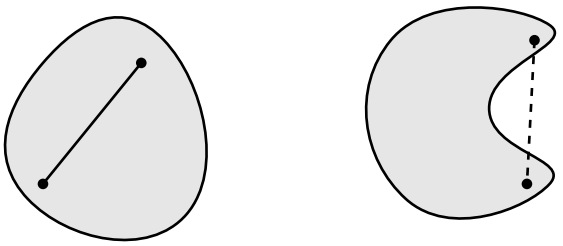


Figure 4.2: Example of a convex set (left) and non-convex set (right).

Definition 4.1.11. *A functional $E\colon \mathcal{X} \to \bar{\mathbb{R}}$ is called* convex*, if*

$$E(\lambda u + (1-\lambda)v) \leqslant \lambda E(u) + (1-\lambda)E(v)$$

for all $\lambda \in (0,1)$ and all $u, v \in \operatorname{dom}(E)$ with $u \neq v$. It is called strictly convex *if the inequality is strict. It is called* strongly *convex with constant θ if $E(u) - \theta\|u\|^2$ is convex.*

Obviously, strong convexity implies strict convexity and strict convexity implies convexity.

Example 4.1.12. The absolute value function $\mathbb{R} \to \mathbb{R}, x \mapsto |x|$ is convex but not strictly convex. The quadratic function $x \mapsto x^2$ is strongly (and hence strictly) convex. The function $x \mapsto x^4$ is strictly convex, but not strongly convex. For other examples, see Figure 4.3.

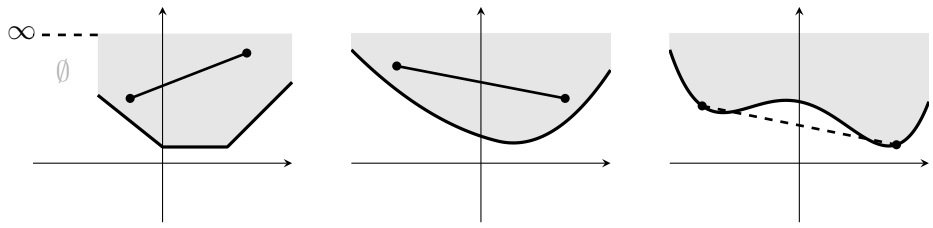


Figure 4.3: Example of a convex function (left), a strictly convex function (middle) and a non-convex function (right).

Example 4.1.13. The characteristic function $\chi_{\mathcal{C}}(u)$ is convex if and only if $\mathcal{C}$ is a convex set. To see the convexity, let $u, v \in \mathrm{dom}(\chi_{\mathcal{C}}) = \mathcal{C}$. Then by the convexity of $\mathcal{C}$ the convex combination $\lambda u + (1-\lambda)v$ is as well in $\mathcal{C}$ and both the left and the right hand side of the desired inequality are zero.

Lemma 4.1.14. *Let $\alpha \geqslant 0$ and $E, F\colon \mathcal{X} \to \bar{\mathbb{R}}$ be two convex functionals. Then $E + \alpha F\colon \mathcal{X} \to \bar{\mathbb{R}}$ is convex. Furthermore, if $\alpha > 0$ and F strictly convex, then $E + \alpha F$ is strictly convex.*

Fenchel conjugate

In convex optimisation problems (i.e. those involving convex functions) the concept of *Fenchel conjugates* plays a very important role.

Definition 4.1.15. *Let $E\colon \mathcal{X} \to \bar{\mathbb{R}}$ be a functional. The functional $E^*\colon \mathcal{X}^* \to \bar{\mathbb{R}}$,*

$$E^*(p) = \sup_{u \in \mathcal{X}} [\langle p, u\rangle - E(u)],$$

is called the Fenchel conjugate of E.

Theorem 4.1.16 ([19, Prop. 4.1])**.** *For any functional $E\colon \mathcal{X} \to \bar{\mathbb{R}}$ the following inequality holds:*

$$E^{**} := (E^*)^* \leqslant E.$$

If E is proper, lower-semicontinuous (see Def. 4.1.3) and convex, then

$$E^{**} = E.$$

Subgradients

For convex functions one can generalise the concept of a derivative so that it would also make sense for non-differentiable functions.

Definition 4.1.17. *A functional $E\colon \mathcal{X} \to \bar{\mathbb{R}}$ is called* subdifferentiable *at $u \in \mathcal{X}$, if there exists an element $p \in \mathcal{X}^*$ such that*

$$E(v) \geqslant E(u) + \langle p, v - u\rangle$$

holds, for all $v \in \mathcal{X}$. Furthermore, we call p a subgradient *at position u. The collection of all subgradients at position u, i.e.*

$$\partial E(u) := \{p \in \mathcal{X}^* \mid E(v) \geqslant E(u) + \langle p, v - u\rangle \ , \forall v \in \mathcal{X}\} \ ,$$

is called subdifferential *of E at u.*

It is clear that if a convex functional $E\colon \mathcal{X} \to \bar{\mathbb{R}}$ is proper, i.e. $\mathrm{dom}(E) \neq \emptyset$, then for all $u \notin \mathrm{dom}(E)$ the subdifferential is empty. A sufficient (but not necessary) condition for E to have a subgradient at $u \in \mathrm{dom}(E)$ is given by

Proposition 4.1.18 ([19, Prop. 5.2])**.** *Let $E\colon \mathcal{X} \to \bar{\mathbb{R}}$ be a convex functional and $u \in \mathrm{dom}(E)$ such that E is continuous at u. Then $\partial E(u) \neq \emptyset$.*

Theorem 4.1.19 ([4, Thm. 7.13])**.** *Let $E\colon \mathcal{X} \to \bar{\mathbb{R}}$ be a proper convex function and $u \in \mathrm{dom}(E)$. Then $\partial E(u)$ is a weak-$*$ compact convex subset of $\mathcal{X}^*$.*

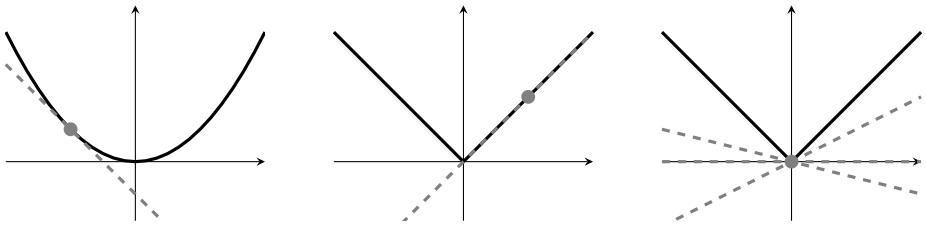


Figure 4.4: Visualisation of the subdifferential. Linear approximations of the functional have to lie completely underneath the function. For points where the function is not differentiable there may be more than one such approximation.

For differentiable functions the subdifferential consists of just one element – the derivative. For non-differentiable functionals the subdifferential is multivalued; we want to consider the subdifferential of the absolute value function as an illustrative example.

Example 4.1.20. Let $E\colon \mathbb{R} \to \mathbb{R}$ be the absolute value function $E(u) = |u|$. Then, the subdifferential of E at u is given by

$$\partial E(u) = \begin{cases} \{1\} & \text{for } u > 0 \\ [-1, 1] & \text{for } u = 0 \\ \{-1\} & \text{for } u < 0 \end{cases},$$

which you will prove as an exercise. A visual explanation is given in Figure 4.4.

The subdifferential of a sum of two functions can be characterised as follows.

Theorem 4.1.21 ([19, Prop. 5.6])**.** *Let $E\colon \mathcal{X} \to \bar{\mathbb{R}}$ and $F\colon \mathcal{X} \to \bar{\mathbb{R}}$ be proper l.s.c. convex functions and suppose $\exists u \in \operatorname{dom}(E) \cap \operatorname{dom}(F)$ such that E is continuous at u. Then*

$$\partial(E + F) = \partial E + \partial F.$$

Using the subdifferential, one can characterise minimisers of convex functionals.

Theorem 4.1.22. *An element $u \in \mathcal{X}$ is a minimiser of the functional $E\colon \mathcal{X} \to \bar{\mathbb{R}}$ if and only if $0 \in \partial E(u)$.*

Proof. By definition, $0 \in \partial E(u)$ if and only if for all $v \in \mathcal{X}$ it holds

$$E(v) \geqslant E(u) + \langle 0, v - u \rangle = E(u)\,,$$

which is by definition the case if and only if u is a minimiser of E. □

The next result connects subgradients and convex conjugates

Theorem 4.1.23 ([19, Prop. 5.1])**.** *Let $E\colon \mathcal{X} \to \bar{\mathbb{R}}$ be a convex function and $E^*\colon \mathcal{X}^* \to \bar{\mathbb{R}}$ its convex conjugate. Then $p \in \partial E(u)$ if and only if*

$$E(u) + E^*(p) = \langle p, u \rangle.$$

Proof. Left as an exercise. □

Bregman distances

Convex functions naturally define some distance measure that became known as the Bregman distance.

Definition 4.1.24. *Let $E\colon \mathcal{X} \to \bar{\mathbb{R}}$ be a convex functional. Moreover, let $u, v \in \mathcal{X}, E(v) < \infty$ and $q \in \partial E(v)$. Then the* (generalised) Bregman distance *of E between u and v is defined as*

$$D_E^q(u, v) := E(u) - E(v) - \langle q, u - v \rangle \ . \tag{4.2}$$

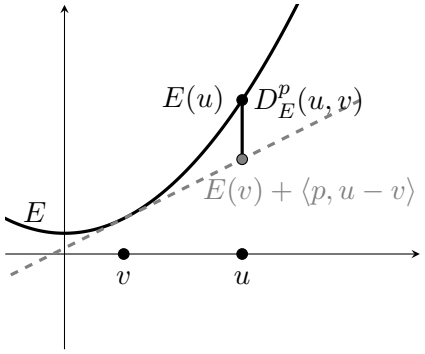


Figure 4.5: Visualization of the Bregman distance.

Remark 4.1.25. It is easy to check that a Bregman distance somewhat resembles a metric as for all $u, v \in \mathcal{X}, q \in \partial E(v)$ we have that $D_E^q(u, v) \geqslant 0$ and $D_E^q(v, v) = 0$. There are functionals where the Bregman distance (up to a square root) is actually a metric; e.g. $E(u) := \frac{1}{2}\|u\|_{\mathcal{X}}^2$ for Hilbert space $\mathcal{X}$, then $D_E^q(u, v) = \frac{1}{2}\|u - v\|_{\mathcal{X}}^2$. However, in general, Bregman distances are not symmetric and $D_E^q(u, v) = 0$ does not imply $u = v$, as you will see on the example sheets.

To overcome the issue of non-symmetry, one can introduce the so-called *symmetric Bregman distance*.

Definition 4.1.26. *Let $E\colon \mathcal{X} \to \bar{\mathbb{R}}$ be a convex functional. Moreover, let $u, v \in \mathcal{X}, E(u) < \infty, E(v) < \infty$, $q \in \partial E(v)$ and $p \in \partial E(u)$. Then the* symmetric Bregman distance *of E between u and v is defined as*

$$D_E^{symm}(u, v) := D_E^q(u, v) + D_E^p(v, u) = \langle p - q, u - v \rangle \ . \tag{4.3}$$

Absolutely one-homogeneous functionals

Definition 4.1.27. *A functional $E\colon \mathcal{X} \to \bar{\mathbb{R}}$ is called absolutely one-homogeneous if*

$$E(\lambda u) = |\lambda| E(u) \quad \forall \lambda \in \mathbb{R},\ \forall u \in \mathcal{X}.$$

Absolutely one-homogeneous convex functionals have some useful properties, for example, it is obvious that $E(0) = 0$. Some further properties are listed below.

Proposition 4.1.28. *Let $E(\cdot)$ be a convex absolutely one-homogeneous functional and let $p \in \partial E(u)$. Then the following equality holds:*

$$E(u) = \langle p, u \rangle.$$

Proof. Left as exercise. □

Remark 4.1.29. The Bregman distance $D_E^p(v,u)$ in this case can be written as follows:

$$D_E^p(v,u) = E(v) - \langle p, v\rangle.$$

Proposition 4.1.30. *Let $E(\cdot)$ be a proper, convex, l.s.c. and absolutely one-homogeneous functional. Then the Fenchel conjugate $E^*(\cdot)$ is the characteristic function of the convex set $\partial E(0)$.*

Proof. Left as exercise. □

An obvious consequence of the above results is the following

Proposition 4.1.31. *For any $u \in \mathcal{X}$, $p \in \partial E(u)$ if and only if $p \in \partial E(0)$ and $E(u) = (p,u)$.*

4.1.3 Minimisers

Definition 4.1.32. *Let $E\colon \mathcal{X} \to \bar{\mathbb{R}}$ be a functional. We say that $u^* \in \mathcal{X}$ solves the minimisation problem*

$$\min_{u\in\mathcal{X}} E(u)$$

if and only if $E(u^*) < \infty$ and $E(u^*) \leqslant E(u)$, for all $u \in \mathcal{X}$. We call u^* a minimiser *of E.*

Definition 4.1.33. *A functional $E\colon \mathcal{X} \to \bar{\mathbb{R}}$ is called* bounded from below *if there exists a constant $C > -\infty$ such that for all $u \in \mathcal{X}$ we have $E(u) \geqslant C$.*

This condition is obviously necessary for the finiteness of the infimum $\inf_{u\in\mathcal{X}} E(u)$.

Existence

If all minimising sequences (that converge to the infimum assuming it exists) are unbounded, then there cannot exist a minimiser. A sufficient condition to avoid such a scenario is *coercivity*.

Definition 4.1.34. *A functional $E\colon \mathcal{X} \to \bar{\mathbb{R}}$ is called* coercive, *if for all $\{u_j\}_{j\in\mathbb{N}}$ with $\|u_j\|_{\mathcal{X}} \to \infty$ we have $E(u_j) \to \infty$.*

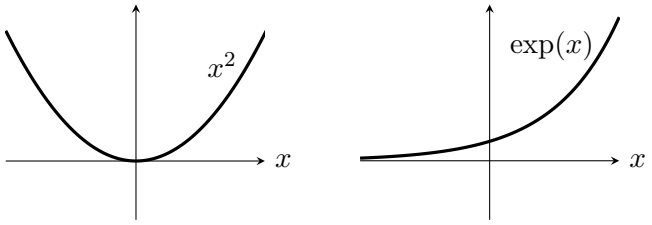


Figure 4.6: While the coercive function on the left has a minimiser, it is easy to see that the non-coercive function on the right does not have a minimiser.

Remark 4.1.35. Coercivity is equivalent to its negated statement which is "if the function values $\{E(u_j)\}_{j\in\mathbb{N}} \subset \mathbb{R}$ are bounded, so is the sequence $\{u_j\}_{j\in\mathbb{N}} \subset \mathcal{X}$".

Although coercivity is not strictly speaking necessary, it is sufficient that all minimising sequences are bounded.

Lemma 4.1.36. *Let $E\colon \mathcal{X} \to \bar{\mathbb{R}}$ be a proper, coercive functional and bounded from below. Then the infimum $\inf_{u\in\mathcal{X}} E(u)$ exists in $\mathbb{R}$, there are minimising sequences, i.e. $\{u_j\}_{j\in\mathbb{N}} \subset \mathcal{X}$ with $E(u_j) \to \inf_{u\in\mathcal{X}} E(u)$, and all minimising sequences are bounded.*

Proof. As E is proper and bounded from below, there exists a $C_1 > 0$ such that we have $-\infty < -C_1 < \inf_u E(u) < \infty$ which also guarantees the existence of a minimising sequence. Let $\{u_j\}_{j\in\mathbb{N}}$ be any minimising sequence, i.e. $E(u_j) \to \inf_u E(u)$. Then there exists a $j_0 \in \mathbb{N}$ such that for all $j > j_0$ we have

$$E(u_j) \leqslant \underbrace{\inf_u E(u) + 1}_{=:C_2} < \infty\,.$$

With $C := \max\{C_1, C_2\}$ we have that $|E(u_j)| < C$ for all $j > j_0$ and thus from the coercivity it follows that $\{u_j\}_{j>j_0}$ is bounded, see Remark 4.1.35. Including a finite number of elements does not change its boundedness which proves the assertion. □

A positive answer about the existence of minimisers is given by the following Theorem known as the "direct method" or "fundamental theorem of optimisation".

Theorem 4.1.37 ("Direct method", David Hilbert, around 1900)**.** *Let $\mathcal{X}$ be a Banach space and $\tau_{\mathcal{X}}$ a topology (not necessarily the one induced by the norm) on $\mathcal{X}$ such that bounded sequences have $\tau_{\mathcal{X}}$-convergent subsequences. Let $E\colon \mathcal{X} \to \bar{\mathbb{R}}$ be proper, bounded from below, coercive and $\tau_{\mathcal{X}}$-l.s.c. Then E has a minimiser.*

Proof. From Lemma 4.1.36 we know that $\inf_{u\in\mathcal{X}} E(u)$ is finite, minimising sequences exist and that they are bounded. Let $\{u_j\}_{j\in\mathbb{N}} \in \mathcal{X}$ be a minimising sequence. Thus, from the assumption on the topology $\tau_{\mathcal{X}}$ there exists a subsequence $\{u_{j_k}\}_{k\in\mathbb{N}}$ and $u^* \in \mathcal{X}$ with $u_{j_k} \overset{\tau_{\mathcal{X}}}{\to} u^*$ for $k \to \infty$. From the sequential lower semi-continuity of E we obtain

$$E(u^*) \leqslant \liminf_{k\to\infty} E(u_{j_k}) = \lim_{j\to\infty} E(u_j) = \inf_{u\in\mathcal{X}} E(u) < \infty\,,$$

which shows that $E(u^*) < \infty$ and $E(u^*) \leqslant E(u)$ for all $u \in \mathcal{X}$; thus u^* minimises E. □

The above theorem is very general but its conditions are hard to verify but the situation is a easier in *reflexive* Banach spaces (thus also in Hilbert spaces).

Corollary 4.1.38. Let $\mathcal{X}$ be a reflexive Banach space and $E\colon \mathcal{X} \to \bar{\mathbb{R}}$ be a functional which is proper, bounded from below, coercive and l.s.c. with respect to the weak topology. Then there exists a minimiser of E.

Proof. The statement follows from the direct method, Theorem 4.1.37, as in reflexive Banach spaces bounded sequences have weakly convergent subsequences, see Theorem 4.1.2. □

Remark 4.1.39. For convex functionals, the situation is even easier. It can be shown that a convex function is l.s.c. with respect to the weak topology if and only if it is l.s.c. with respect to the strong topology (see e.g. [19, Corollary 2.2., p. 11] or [8, p. 149] for Hilbert spaces).

Remark 4.1.40. It is easy to see that the key ingredient for the existence of minimisers is that bounded sequences have a convergent subsequence. In variational regularisation this is usually ensured by an appropriate choice of the regularisation functional.

Uniqueness

Theorem 4.1.41. *Assume that the functional $E\colon \mathcal{X} \to \bar{\mathbb{R}}$ has at least one minimiser and is strictly convex. Then the minimiser is unique.*

Proof. Let u, v be two minimisers of E and assume that they are different, i.e. $u \neq v$. Then it follows from the minimising properties of u and v as well as the strict convexity of E that

$$E(u) \leqslant E(\tfrac{1}{2}u + \tfrac{1}{2}v) < \frac{1}{2}E(u) + \frac{1}{2}\underbrace{E(v)}_{\leqslant E(u)} \leqslant E(u)$$

which is a contradiction. Thus, $u = v$ and the assertion is proven. □

Example 4.1.42. Convex (but not strictly convex) functions may have have more than one minimiser, examples include constant and trapezoidal functions, see Figure 4.7. On the other hand, convex (and even non-convex) functions may have a unique minimiser, see Figure 4.7.

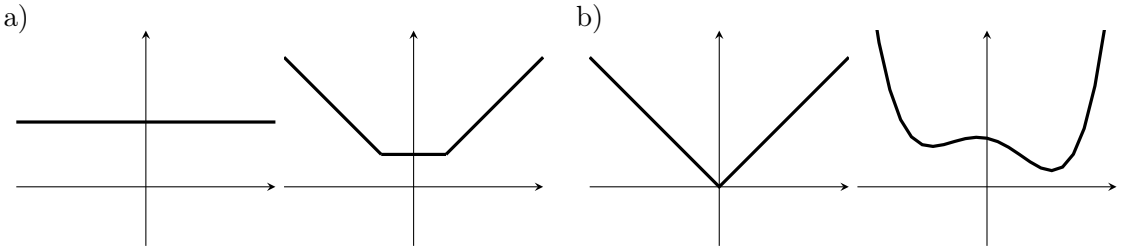


Figure 4.7: a) Convex functions may not have a unique minimiser. b) Neither strict convexity nor convexity is necessary for the uniqueness of a minimiser.

4.1.4 Duality in convex optimisation

Consider the following optimisation problem

$$\inf_{u \in \mathcal{X}} E(Au) + F(u), \tag{$\mathcal{P}$}$$

where $E\colon \mathcal{Y} \to \bar{\mathbb{R}}$ and $F\colon \mathcal{X} \to \bar{\mathbb{R}}$ are proper, convex and lower semicontinuous functions and $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ is a linear bounded operator.

Since E is convex and lower semicontinuous, it can be written as the convex conjugate of its conjugate E^*

$$E(y) = \sup_{\eta \in \mathcal{Y}^*} \langle \eta, y \rangle - E^*(\eta) \quad y \in \mathcal{Y}.$$

Hence, we can rewrite ($\mathcal{P}$) as follows

$$\inf_{u \in \mathcal{X}} \sup_{\eta \in \mathcal{Y}^*} \langle \eta, Au \rangle - E^*(\eta) + F(u). \tag{$\mathcal{S}$}$$

This problem is referred to as the *saddle point problem*, whereas $(\mathcal{P})$ is referred to as the *primal problem*. Since $\inf\sup \geqslant \sup\inf$ always holds, we get that

$$\begin{aligned}\inf_{u\in\mathcal{X}} E(Au)+F(u) &\geqslant \sup_{\eta\in\mathcal{Y}^*}\inf_{u\in\mathcal{X}}\langle\eta,Au\rangle - E^*(\eta)+F(u)\\ &= \sup_{\eta\in\mathcal{Y}^*}\inf_{u\in\mathcal{X}}\langle A^*\eta,u\rangle - E^*(\eta)+F(u)\\ &= \sup_{\eta\in\mathcal{Y}^*}\left\{-E^*(\eta)-\sup_{u\in\mathcal{X}}\left[\langle -A^*\eta,u\rangle - F(u)\right]\right\}\\ &= \sup_{\eta\in\mathcal{Y}^*} -E^*(\eta)-F^*(-A^*\eta).\end{aligned}$$

The last problem

$$\sup_{\eta\in\mathcal{Y}^*} -E^*(\eta)-F^*(-A^*\eta) \tag{$\mathcal{D}$}$$

is called the *dual problem*. The fact that the optimal value of the primal is always less or equal to the optimal value of the dual problem is referred to as *weak duality* and the difference between these two optimal values is referred to as the *duality gap*. Whenever the two optimal values are in fact equal, one speaks of *strong duality*. Sufficient conditions for strong duality are given by

Theorem 4.1.43 ([19, Ch.III Thm 4.1 and Rem. 4.2])**.** *Suppose that*

(i) the function $E(Au)+F(u)\colon \mathcal{X}\to\bar{\mathbb{R}}$ is proper, convex, l.s.c. and coercive;

(ii) $\exists u_0\in\mathcal{X}$ s.t. $F(u_0)<+\infty$, $E(Au_0)<+\infty$ and $E(y)$ is continuous at $y=Au_0$.

Then

(i) The dual problem $(\mathcal{D})$ has at least one solution $\widehat{\eta}$;

(ii) There is no duality gap between $(\mathcal{P})$ and $(\mathcal{D})$, i.e. strong duality holds;

(iii) If $(\mathcal{P})$ has an optimal solution $\widehat{u}$, then the following optimality conditions hold

$$-A^*\widehat{\eta}\in\partial F(\widehat{u}),\quad \widehat{\eta}\in\partial E(A\widehat{u}).$$

Note that existence of a primal solution is *not* guaranteed by this theorem.

4.2 Well-posedness and Regularisation Properties

Our goal is to study the properties of optimisation problem (4.1) as a convergent regularisation for the ill-posed problem

$$Au=f, \tag{4.4}$$

where $A\colon\mathcal{X}\to\mathcal{Y}$ is a linear bounded operator, $\mathcal{Y}$ is a Banach space and $\mathcal{X}$ is the dual of a separable Banach space. In particular, we will ask questions of existence of minimisers (well-posedness of the regularised problem) and parameter choice rules that guarantee the convergence of the minimisers to an appropriate generalised solution of (4.4) for different choices of the regularisation functional. To this end, we need to extend the definition of a minimal-norm solution (Def. 2.1.1) to an arbitrary regularisation term.

Definition 4.2.1 ($\mathcal{J}$-minimising solutions)**.** *Let $u^\dagger_{\mathcal{J}}$ be a least squares solution, i.e.*

$$\|Au^\dagger_{\mathcal{J}} - f\|_{\mathcal{Y}} = \inf\{\|Av - f\|_{\mathcal{Y}}, \quad v \in \mathcal{X}\}$$

and

$$\mathcal{J}(u^\dagger_{\mathcal{J}}) \leqslant \mathcal{J}(\widetilde{u}) \quad \textit{for all least squares solutions } \widetilde{u}.$$

Then $u^\dagger_{\mathcal{J}}$ is called a $\mathcal{J}$-minimising solution of (4.4).

We will assume that there exists a least-squares solution with a finite value of $\mathcal{J}$, i.e. there exists at least one element u such that $\|Au - f\|_{\mathcal{Y}} = \inf\{\|Av - f\|_{\mathcal{Y}}, v \in \mathcal{X}\}$ and $\mathcal{J}(u) < +\infty$.

Remark 4.2.2. A $\mathcal{J}$-minimising solution may not exist and if it does, it may be non-unique. We will later see conditions, under which a $\mathcal{J}$-minimising solution exists. Non-uniqueness, however, is common with popular choices of $\mathcal{J}$. In this case we need to define a *selection operator* that will select a single element from all the $\mathcal{J}$-minimising solutions (see [9]). We will not explicitly mention this, stating all results for just *a* $\mathcal{J}$-minimising solution.

We will need the following

Lemma 4.2.3. *Let $\mathcal{J}(u) = \sum_{i=1}^n \mathcal{J}_i(u)$, where each $\mathcal{J}_i(u)$ is convex and p_i-homogeneous ($p_i > 0$), that is,*

$$\mathcal{J}_i(\lambda u) = |\lambda|^{p_i}\mathcal{J}_i(u) \quad \forall u \in \mathcal{X}, \lambda \in \mathbb{R}.$$

The the set

$$\mathcal{N}(\mathcal{J}) := \{u \in \mathcal{X} \colon \mathcal{J}(u) = 0\}$$

is a linear subspace of $\mathcal{X}$.

Proof. First of all, we note that $\mathcal{J}_i(u) \geqslant 0$ for all $u \in \mathcal{X}$. Indeed, we have

$$0 = \mathcal{J}_i(0) = \mathcal{J}_i\left(\frac{1}{2}u - \frac{1}{2}u\right) \leqslant \frac{1}{2}\mathcal{J}_i(u) + \frac{1}{2}\mathcal{J}_i(-u) = \mathcal{J}_i(u).$$

Now let $u, v \in \mathcal{N}(\mathcal{J})$ be arbitrary. Then $\mathcal{J}_i(u) = \mathcal{J}_i(v) = 0$ for all $i = 1, ..., n$, hence for any $\lambda \in \mathbb{R}$

$$\begin{aligned} 0 \leqslant \mathcal{J}_i(\lambda u + v) &= 2^{p_i}\mathcal{J}_i\left(\frac{\lambda u}{2} + \frac{v}{2}\right) \leqslant 2^{p_i}\left(\frac{1}{2}\mathcal{J}_i\left(\frac{\lambda u}{2}\right) + \frac{1}{2}\mathcal{J}_i\left(\frac{v}{2}\right)\right) \\ &= \frac{1}{2}\mathcal{J}_i(\lambda u) + \frac{1}{2}\mathcal{J}_i(v) = \frac{|\lambda|^{p_i}}{2}\mathcal{J}_i(u) + \frac{1}{2}\mathcal{J}_i(v) = 0. \end{aligned}$$

Therefore, $\mathcal{J}_i(\lambda u + v) = 0$ for all i and hence $\mathcal{J}(\lambda u + v) = 0$. □

Lemma 4.2.4. *Let assumptions of Lemma 4.2.3 be satisfied. Suppose that $u \in \mathcal{X}$ and $v \in \mathcal{N}(\mathcal{J})$. Then $\mathcal{J}(u + v) = \mathcal{J}(u)$.*

Proof. Left as exercise. □

If $\dim \mathcal{N}(\mathcal{J}) < \infty$, the subspace $\mathcal{N}(\mathcal{J})$ is *complemented* in $\mathcal{X}$ [4, Thm. 5.89], i.e. there exists a closed subspace $\mathcal{X}_0 \subset \mathcal{X}$ such that $\mathcal{X}_0 \cap \mathcal{N}(\mathcal{J}) = \{0\}$ and

$$\mathcal{X} = \mathcal{X}_0 \oplus \mathcal{N}(\mathcal{J}). \tag{4.5}$$

We will use this to establish coercivity of the functional (4.1).

Lemma 4.2.5. *Suppose that the regularisation functional $\mathcal{J}\colon \mathcal{X} \to \bar{\mathbb{R}}_+$ is proper, convex and satisfies conditions of Lemma* (4.2.3) *and let $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ be a bounded linear operator. Suppose also that*

(i) $\dim \mathcal{N}(\mathcal{J}) < \infty$ *and $\mathcal{J}$ is coercive on $\mathcal{X}_0$, where $\mathcal{X}_0$ is such that $\mathcal{X} = \mathcal{X}_0 \oplus \mathcal{N}(\mathcal{J})$;*

(ii) the kernels of A and $\mathcal{J}$ have a trivial intersection, i.e. $\mathcal{N}(A) \cap \mathcal{N}(\mathcal{J}) = \{0\}$.

Then the function

$$\Phi_\alpha(u) := \frac{1}{2}\|Au - f\|_{\mathcal{Y}}^2 + \alpha \mathcal{J}(u)$$

is coercive on $\mathcal{X}$ for any $\alpha > 0$.

Proof. Let $\{u_j\}_{j\in\mathbb{N}}$ be a sequence in $\mathcal{X}$. Due to (4.5), there exists a unique decomposition

$$u_j = u_j^0 + u_j^{\mathcal{N}}, \quad u_j^0 \in \mathcal{X}_0,\ u_j^{\mathcal{N}} \in \mathcal{N}(\mathcal{J}).$$

Let $\Phi_\alpha(u_j) \leqslant C$ for all $j \in \mathbb{N}$. Then $\mathcal{J}(u_j) \leqslant C$ and

$$\mathcal{J}\left(u_j^0\right) = \mathcal{J}\left(u_j^0 + u_j^{\mathcal{N}}\right) = \mathcal{J}(u_j) \leqslant C.$$

Since $\mathcal{J}$ is coercive on $\mathcal{X}_0$, we get that $\|u_j^0\| \leqslant C'$. Now, define

$$\widetilde{A}\colon \mathcal{N}(\mathcal{J}) \to A\mathcal{N}(\mathcal{J}), \quad \widetilde{A} = A|_{\mathcal{N}(\mathcal{J})}.$$

That is, $\widetilde{A}$ is the restriction of A to $\mathcal{N}(\mathcal{J})$. Clearly, $\widetilde{A}$ is surjective and by assumption (ii) it is also injective. Since $\mathcal{N}(\mathcal{J})$ (and, subsequently, $A\mathcal{N}(\mathcal{J})$) is finite-dimensional, $\widetilde{A}^{-1}$ exists and is bounded. Denote $\|\widetilde{A}^{-1}\| =: \widetilde{C}$. Then

$$\begin{aligned} \|u_j^{\mathcal{N}}\| &= \|\widetilde{A}^{-1}(\widetilde{A}u_j^{\mathcal{N}})\| \leqslant \widetilde{C}\|Au_j^{\mathcal{N}}\| = \widetilde{C}\|Au_j^{\mathcal{N}} + Au_j^0 - f - (Au_j^0 - f)\| \\ &\leqslant \widetilde{C}\left\|Au_j - f\| + \|Au_j^0 - f\|\right) \leqslant \widetilde{C}(C + \|A\|\|u_j^0\| + \|f\|) \leqslant C''. \end{aligned}$$

Therefore,

$$\|u_j\| = \|u_j^0 + u_j^{\mathcal{N}}\| \leqslant \|u_j^0\| + \|u_j^{\mathcal{N}}\| \leqslant C''',$$

which means that Φ_α is coercive. □

Now we are ready to establish the existence of a $\mathcal{J}$-minimising solution and a regularised solution for any $\alpha > 0$.

Theorem 4.2.6. *Let $\mathcal{X}$ and $\mathcal{Y}$ be a Banach spaces and $\tau_{\mathcal{X}}$ and $\tau_{\mathcal{Y}}$ some topologies (not necessarily induced by the norm) in $\mathcal{X}$ and $\mathcal{Y}$, respectively. Assume that*

(i) bounded sequences in $\mathcal{X}$ have $\tau_{\mathcal{X}}$-convergent subsequences;

(ii) $\mathcal{J}\colon \mathcal{X} \to \bar{\mathbb{R}}_+$ is proper, convex $\tau_{\mathcal{X}}$-l.s.c. and satisfies assumptions of Lemma 4.2.5;

(iii) $A\colon \mathcal{X} \to \mathcal{Y}$ is $\tau_{\mathcal{X}} \to \tau_{\mathcal{Y}}$ continuous;

(iv) $\|\cdot\|_{\mathcal{Y}}$ is $\tau_{\mathcal{Y}}$-lower semicontinuous;

Then

(i') there exists a $\mathcal{J}$-minimising solution $u_{\mathcal{J}}^\dagger$ of (4.4)*;*

(ii') for any fixed $\alpha > 0$ and $f \in \mathcal{Y}$ there exists a minimiser

$$u^\alpha \in \operatorname*{arg\,min}_{u\in\mathcal{X}} \frac{1}{2}\|Au - f\|_{\mathcal{Y}}^2 + \alpha\mathcal{J}(u). \tag{4.6}$$

Proof. (i) Let $\mathbb{L}$ be the set of least-squares solutions of (4.4). Then $\mathbb{L}$ can written as follows

$$\mathbb{L} = \{u \in \mathcal{X} \colon \|Au - f\|_{\mathcal{Y}} \leqslant \mu\},$$

where $\mu := \inf\{\|Av - f\|_{\mathcal{Y}} \colon v \in \mathcal{X}\}$. Since A is $\tau_{\mathcal{X}} \to \tau_{\mathcal{Y}}$ continuous and $\|\cdot\|_{\mathcal{Y}}$ is $\tau_{\mathcal{Y}}$-l.s.c., $\mathbb{L}$ is $\tau_{\mathcal{X}}$-closed.

Consider the following problem

$$\inf_{u\in\mathbb{L}} \mathcal{J}(u) = \inf_{u\in\mathcal{X}} \mathcal{J}(u) + \chi_{\mathbb{L}}(u). \tag{4.7}$$

By the assumption that we made in the beginning of this section, this problem is feasible, i.e. there exists $u \in \mathbb{L}$ with $\mathcal{J}(u) < \infty$. The objective function in (4.7) is bounded from below. Using similar arguments as in Lemma 4.2.5, we conclude that it is also coercive. Since $\mathbb{L}$ is $\tau_{\mathcal{X}}$-closed, $\chi\mathbb{L}$ is $\tau_{\mathcal{X}}$-l.s.c. By assumption ii, $\mathcal{J}$ is also $\tau_{\mathcal{X}}$-l.s.c. So, (4.7) satisfies the assumptions of the direct method (Theorem 4.1.37) and hence a minimiser exists.

(ii) From Lemma 4.2.5 we know that the objective function Φ_α in (4.6) is coercive. It is also bounded from below. Since $\mathcal{J}$ is $\tau_{\mathcal{X}}$-l.s.c., A is $\tau_{\mathcal{X}} \to \tau_{\mathcal{Y}}$ continuous and $\|\cdot\|_{\mathcal{Y}}$ is $\tau_{\mathcal{Y}}$-l.s.c., we get that Φ_α is $\tau_{\mathcal{X}}$-l.s.c. Using the direct method, we conclude that (4.6) has a minimiser.

□

Now we study the behaviour of the minimiser of (4.6) with $f = f_\delta$ (perturbed measurement) as $\delta \to 0$ when $\alpha = \alpha(\delta)$ is chosen according to an appropriate a priori parameter choice rule. For simplicity, we will do this in the case when $\inf\{\|Av - f\|_{\mathcal{Y}} \colon v \in \mathcal{X}\} = 0$, i.e. least-squares solutions are actually solutions of (4.4).

Theorem 4.2.7. *Let the assumptions of Theorem 4.2.6 hold and suppose that* $\inf\{\|Av - f\|_{\mathcal{Y}} \colon v \in \mathcal{X}\} = 0$. *Let $\alpha = \alpha(\delta)$ be such that*

$$\lim_{\delta\to 0} \alpha(\delta) = 0 \quad \textit{and} \quad \limsup_{\delta\to 0} \frac{\delta^2}{\alpha(\delta)} = 0.$$

Then $u_\delta := u_\delta^{\alpha(\delta)} \overset{\tau_{\mathcal{X}}}{\to} u_{\mathcal{J}}^\dagger$ as $\delta \to 0$ (possibly, along a subsequence) and $\mathcal{J}(u_\delta) \to \mathcal{J}(u_{\mathcal{J}}^\dagger)$, where $u_{\mathcal{J}}^\dagger$ is a $\mathcal{J}$-minimising solution.

Proof. Let u_0 be any $\mathcal{J}$-minimising solution (which exists by Theorem 4.2.6). Since u_δ solves (4.6) with $\alpha = \alpha(\delta)$, we get that

$$\begin{aligned} \frac{1}{2}\|Au_\delta - f_\delta\|_{\mathcal{Y}}^2 + \alpha(\delta)\mathcal{J}(u_\delta) &\leqslant \frac{1}{2}\|Au_0 - f_\delta\|_{\mathcal{Y}}^2 + \alpha(\delta)\mathcal{J}(u_0) \\ &\leqslant \frac{\delta^2}{2} + \alpha(\delta)\mathcal{J}(u_0). \end{aligned} \tag{4.8}$$

Therefore, we have the following two estimates

$$\mathcal{J}(u_\delta) \leqslant \frac{\delta^2}{2\alpha(\delta)} + \mathcal{J}(u_0) \leqslant C, \tag{4.9a}$$

$$\|Au_\delta - f_\delta\|_{\mathcal{Y}} \leqslant \sqrt{\delta^2 + 2\alpha(\delta)\mathcal{J}(u_0)} \leqslant C', \tag{4.9b}$$

The right-hand side in (4.9a) is bounded uniformly in δ, because $\limsup_{\delta\to 0} \delta^2/\alpha(\delta) = 0$ by assumption and $\mathcal{J}(u_0)$ is a constant independent of δ. The right-hand side in (4.9b) is bounded, because $\mathcal{J}(u_0)$ is a constant and $\delta, \alpha(\delta) \to 0$.

Therefore, both $\mathcal{J}(u_\delta)$ and $\|Au_\delta - f_\delta\|_{\mathcal{Y}}$ are uniformly bounded. Proceeding similarly to Lemma 4.2.5, we get that

$$\|u_\delta\| \leqslant C$$

for all δ. Now let $\delta_n \downarrow 0$ be an arbitrary null sequence. Since u_{δ_n} is bounded, it contains a $\tau_{\mathcal{X}}$-convergent subsequence (which we don't relabel)

$$u_{\delta_n} \overset{\tau_{\mathcal{X}}}{\to} u_{\mathcal{J}}^\dagger \quad \text{as } n \to \infty.$$

We will show that $u_{\mathcal{J}}^\dagger$ is a $\mathcal{J}$-minimising solution. From (4.9b) we observe that

$$\liminf_{n\to\infty} \|Au_{\delta_n} - f_{\delta_n}\|_{\mathcal{Y}} \leqslant \liminf_{n\to\infty} \sqrt{\delta_n^2 + 2\alpha(\delta_n)\mathcal{J}(u_0)} = 0.$$

Since A is $\tau_{\mathcal{X}} \to \tau_{\mathcal{Y}}$ continuous and $\|\cdot\|_{\mathcal{Y}}$ is $\tau_{\mathcal{Y}}$-l.s.c., we get that

$$\|Au_{\mathcal{J}}^\dagger - f\|_{\mathcal{Y}} \leqslant \liminf_{n\to\infty} \|Au_{\delta_n} - f\|_{\mathcal{Y}} \leqslant \liminf_{n\to\infty} \left(\|Au_{\delta_n} - f_{\delta_n}\|_{\mathcal{Y}} + \|f - f_{\delta_n}\|_{\mathcal{Y}}\right) = 0,$$

which shows that $u_{\mathcal{J}}^\dagger$ is a least-squares solution. Using the estimate (4.9a) and $\tau_{\mathcal{X}}$-lower semicontinuity of $\mathcal{J}$, we obtain

$$\mathcal{J}(u_{\mathcal{J}}^\dagger) \leqslant \liminf_{n\to\infty} \mathcal{J}(u_{\delta_n}) \leqslant \limsup_{n\to\infty} \mathcal{J}(u_{\delta_n}) \leqslant \limsup_{n\to\infty} \frac{\delta^2}{2\alpha(\delta)} + \mathcal{J}(u_0) = \mathcal{J}(u_0). \tag{4.10}$$

Since u_0 was an arbitrary $\mathcal{J}$-minimising solution and $\mathcal{J}(u_{\mathcal{J}}^\dagger) \leqslant \mathcal{J}(u_0)$, we conclude that $\mathcal{J}(u_{\mathcal{J}}^\dagger)$ is also a $\mathcal{J}$-minimising solution. Finally, since $\mathcal{J}(u_{\mathcal{J}}^\dagger) = \mathcal{J}(u_0)$, we conclude from (4.10) that

$$\liminf_{n\to\infty} \mathcal{J}(u_{\delta_n}) = \limsup_{n\to\infty} \mathcal{J}(u_{\delta_n}) = \lim_{n\to\infty} \mathcal{J}(u_{\delta_n}) = \mathcal{J}(u_{\mathcal{J}}^\dagger),$$

which completes the proof. □

Remark 4.2.8. The theorem proves convergence of the regularised solutions in $\tau_{\mathcal{X}}$, which may differ from the strong topology. However, if $\mathcal{J}$ satisfies the *Radon-Riesz property* with respect to the topology $\tau_{\mathcal{X}}$, i.e. $u_j \overset{\tau_{\mathcal{X}}}{\to} u$ and $\mathcal{J}(u_j) \to \mathcal{J}(u)$ imply $\|u_j - u\| \to 0$, then we get convergence in the norm topology. An example of a functional satisfying the Radon-Riesz property is the norm in a Hilbert (or reflexive Banach) space with $\tau_{\mathcal{X}}$ being the weak topology.

Examples of regularisers

Example 4.2.9. Let $\mathcal{X}$ be a Hilbert space and $\mathcal{J}(u) = \|u\|^2$. The norm in a Hilbert space is weakly l.s.c. By Theorem 4.1.2 we know that (norm) bounded sequences have weakly convergent subsequences. Therefore, Assumption (ii) of Theorem 4.2.6 is satisfied with $\tau_{\mathcal{X}}$ being the weak topology and we obtain weak convergence of the regularised solutions. However, since the norm in a Hilbert space has the Radon-Riesz property, we also get strong convergence. The same approach works in reflexive Banach spaces.

A classical example is regularisation in Sobolev spaces such as the space H^1 of L^2 functions whose weak derivatives are also in L^2. In the one-dimensional case, the space H^1 consists only of continuous functions (in higher dimensions it is true for Sobolev spaces with some other exponents), therefore, the regularised solutions will also be continuous. For this reason, the regulariser $\mathcal{J}(u) = \|u\|_{H^1}$ is sometimes referred to as the *smoothing functional*. Whilst desirable in some applications, in imaging smooth reconstructions are usually not favourable, since images naturally contain edges and therefore are not continuous functions. To overcome this issue, other regularisers have been introduced that we will discuss later.

Example 4.2.10 (ℓ^1-regularisation)**.** Let $\mathcal{X} = \ell^2$ be space of all square summable sequences (i.e. such that $\|u\|_{\ell^2}^2 = \sum_{i=1}^{\infty} u_i^2 < +\infty$). For example, u can represent the coefficients of a function in a basis (e.g., a Fourier basis or a wavelet basis). As a regularisation functional, let us use not the ℓ^2-norm, but the ℓ^1-norm:

$$\mathcal{J}(u) = \|u\|_{\ell^1} = \sum_{i=1}^{\infty} |u_i|.$$

By Example 4.1.5 $\mathcal{J}(\cdot)$ is weakly l.s.c. in ℓ^2. It is evident that $\ell^q \subset \ell^p$ and $\|\cdot\|_{\ell^p} \leqslant \|\cdot\|_{\ell^q}$ for $q \leqslant p$. Therefore, $\mathcal{J}(u) \leqslant C$ implies that $\|\cdot\|_{\ell^2} \leqslant C$ and, since ℓ^2 is a Hilbert space and bounded sequences have weakly convergent subsequences, we conclude that the sublevel sets of $\mathcal{J}(\cdot)$ are weakly sequentially compact in ℓ^2. Therefore, Assumption (ii) of Theorem 4.2.6 is satisfied with $\tau_{\mathcal{X}}$ being the weak topology in ℓ^2. Hence, we get weak convergence of regularised solutions in ℓ^2.

The motivation for using the ℓ^1-norm as the regulariser instead of the ℓ^2-norm is as follows. If the forward operator is non-injective, the inverse problem has more than one solution and the solutions form an affine subspace. In the context of sequence spaces representing coefficients of the solution in a basis, it is sometimes beneficial to look for solutions that are *sparse* in the sense that they have finite support, i.e. $|\operatorname{supp}(u)| < \infty$ with $\operatorname{supp}(u) = \{i \in \mathbb{N} \,|\, u_i \neq 0\}$. This allows explaining the signal with a finite (and often relatively small) number of basis functions and has widely ranging applications in, for instance, compressed sensing. A finite dimensional illustration of the sparsity of ℓ^1-regularised solutions is given in Figure 4.8. The corresponding minimisation problem

$$\min_{u \in \ell^2} \left\{ \frac{1}{2} \|Au - f\|_{\ell^2}^2 + \alpha \|u\|_1 \right\}. \tag{4.11}$$

is also called *lasso* in the statistical literature.

Example 4.2.11 (Elastic net regularisation)**.** The ℓ^1 regulariser described in the previous example sometimes delivers undesirable results for problems where there are highly correlated features and we need to identify all relevant ones, e.g. microarray data analysis

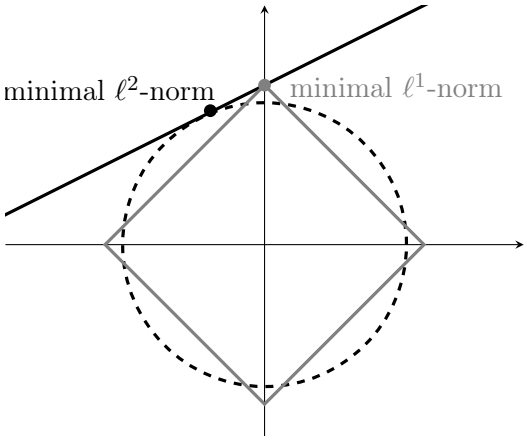


Figure 4.8: Non-injective operators have a non-trivial kernel such that the inverse problem has more than one solution and the solutions form an affine subspace visualised by the solid line. Different regularisation functionals favour different solutions. The circle and the diamond indicate all points with constant ℓ^2-norm, respectively ℓ^1-norm, and the minimal ℓ^2-norm and ℓ^1-norm solutions are the intersections of the line with the circle, respectively the diamond. As it can be seen, the minimal ℓ^2-norm solution has two non-zero components while the minimal ℓ^1-norm solution has only one non-zero component and thus is *sparser*.

(analysis of genomic sequences), in that it tends to select only one feature out of the relevant group instead of all relevant features of the group, i.e. it fails to identify the group structure. Elastic net regularisation helps to overcome this issue. The elastic net regulariser $\mathcal{J}\colon \ell^2 \to \bar{\mathbb{R}}_+$ is defined as follows

$$\mathcal{J}(u) := \alpha\|u\|_{\ell^1} + \beta\|u\|_{\ell^2}^2,$$

where $\alpha, \beta > 0$ are constants that balance the influence of the two terms. Since $\mathcal{J}$ is the sum of a 1-homogeneous term and a 2-homogeneous term, it satisfies assumptions of Lemma 4.2.3.

4.3 Total Variation Regularisation

As pointed out in Example 4.2.9, in imaging we are interested in regularisers that allow for discontinuities while maintaining sufficient regularity of the reconstructions. One popular choice is the so-called *total variation* regulariser [15].

Definition 4.3.1. *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain and $u \in L^1(\Omega)$. Let $\mathcal{D}(\Omega, \mathbb{R}^n)$ be the following set of vector-valued* test functions *(i.e. functions that map from Ω to $\mathbb{R}^n$)*

$$\mathcal{D}(\Omega, \mathbb{R}^n) := \Big\{\varphi \in C_0^\infty(\Omega; \mathbb{R}^n) \;\Big|\; \sup_{x\in\Omega} \|\varphi(x)\|_2 \leqslant 1\Big\}.$$

Total variation of $u \in L^1(\Omega)$ is defined as follows

$$\mathrm{TV}(u) = \sup_{\varphi\in\mathcal{D}(\Omega,\mathbb{R}^n)} \int_\Omega u(x) \operatorname{div} \varphi(x)\, dx\,.$$

Remark 4.3.2. Definition 4.3.1 may seem a bit strange at the first glance, but we note that for a function $u \in L^1(\Omega)$ whose weak derivative ∇u exists and is also in $L^1(\Omega, \mathbb{R}^n)$ (i.e. u belongs to the Sobolev space $W^{1,1}(\Omega)$) we obtain, integrating by parts, that

$$\mathrm{TV}(u) = \sup_{\varphi \in \mathcal{D}(\Omega, \mathbb{R}^n)} \int_\Omega - \langle \nabla u(x), \varphi(x) \rangle \, dx.$$

By the Cauchy-Schwartz inequality we get that $|\langle \nabla u(x), \varphi(x) \rangle| \leqslant \|\nabla u(x)\|_2 \|\varphi(x)\|_2 \leqslant \|\nabla u(x)\|_2$ for a.e. $x \in \Omega$. On the other hand, choosing φ such that $\varphi(x) = -\frac{\nabla u(x)}{\|\nabla u(x)\|_2}$ (technically, such φ is not necessarily in $\mathcal{D}(\Omega, \mathbb{R}^n)$, but we can approximate it with functions from $\mathcal{D}(\Omega, \mathbb{R}^n)$, since any function in $W^{1,1}(\Omega)$ can be approximated with smooth functions [2, Thm. 3.17]; we omit the technicalities here), we get that $-\langle \nabla u(x), \varphi(x) \rangle = \|\nabla u(x)\|_2$. Therefore, the supremum over $\varphi \in \mathcal{D}(\Omega, \mathbb{R}^n)$ is equal to

$$\mathrm{TV}(u) = \int_\Omega \|\nabla u(x)\|_2 \, dx = \|\nabla u\|_{L^1}.$$

This shows that TV just penalises the the L^1 norm (of the pointwise 2-norm) of the gradient for any $u \in W^{1,1}(\Omega)$. However, we will see that the space of functions that have finite value of TV is larger than $W^{1,1}(\Omega)$ and contains, for instance, discontinuous functions.

Remark 4.3.3. It can be shown [13] that for any $u \in L^1(\Omega)$

$$\mathrm{TV}(u) = \|\nabla u\|_{\mathfrak{M}},$$

where ∇ is the distributional gradient and $\|\cdot\|_{\mathfrak{M}}$ is the Radon norm. That is, Total Variation extends the L^1 norm of the gradient for functions whose gradient is not a Lebesgue-measurable function. We will not use this interpretation of the Total Variation to simplify the presentation and refer the interested reader to [13] for details.

Proposition 4.3.4. TV *is a proper, convex and absolutely* 1*-homogeneous functional* $L^1(\Omega) \to \bar{\mathbb{R}}$. *For any constant function* $\mathbf{c}\colon \mathbf{c}(x) \equiv c \in \mathbb{R}$ *for all* x *and any* $u \in L^1(\Omega)$

$$\mathrm{TV}(\mathbf{c}) = 0 \quad \textit{and} \quad \mathrm{TV}(u + \mathbf{c}) = \mathrm{TV}(u).$$

Proof. Left as exercise. □

Remark 4.3.5. It can be shown that the opposite implication holds, i.e. $\mathrm{TV}(u) = 0$ implies that u is constant. in other words,

$$\mathcal{N}(\mathrm{TV}) = \{u \in L^1(\Omega)\colon u = const\}. \tag{4.12}$$

The easiest way to see this is using the Radon measure interpretation in Remark 4.3.3. Because time constraints, we will omit the proof.

Example 4.3.6 (TV of an indicator function)**.** Suppose $\mathcal{C} \subset \Omega \subset \mathbb{R}^2$ is a bounded domain with smooth boundary and $u(\cdot) = \mathbf{1}_{\mathcal{C}}(\cdot)$ is its indicator function, i.e.

$$\mathbf{1}_{\mathcal{C}}(x) = \begin{cases} 1 & x \in \mathcal{C} \\ 0 & x \in \mathcal{X} \setminus \mathcal{C} \end{cases}.$$

Then, using the divergence theorem, we get that for any test function $\varphi \in \mathcal{D}(\Omega, \mathbb{R}^n)$

$$\int_\Omega u(x) \operatorname{div} \varphi(x)\, dx = \int_{\mathcal{C}} \operatorname{div} \varphi(x)\, dx = \int_{\partial \mathcal{C}} \langle \varphi(x), \mathbf{n}_{\partial \mathcal{C}}(x) \rangle \; dl,$$

where $\partial \mathcal{C}$ is the boundary of $\mathcal{C}$ and $\mathbf{n}_{\partial \mathcal{C}}(x)$ is the unit normal at x. Hence,

$$\begin{aligned} \mathrm{TV}(u) &= \sup_{\varphi \in \mathcal{D}(\Omega, \mathbb{R}^n)} \int_\Omega u(x) \operatorname{div} \varphi(x)\, dx = \sup_{\varphi \in \mathcal{D}(\Omega, \mathbb{R}^n)} \int_{\partial \mathcal{C}} \langle \varphi(x), \mathbf{n}_{\partial \mathcal{C}}(x) \rangle \; dl \\ &\leqslant \sup_{\varphi \in \mathcal{D}(\Omega, \mathbb{R}^n)} \int_{\partial \mathcal{C}} \|\varphi(x)\| \|\mathbf{n}_{\partial \mathcal{C}}(x)\| \, dl \leqslant \sup_{\varphi \in \mathcal{D}(\Omega, \mathbb{R}^n)} \int_{\partial \mathcal{C}} dl = \mathrm{Per}_{\mathcal{C}}, \end{aligned}$$

where $\mathrm{Per}(\mathcal{C})$ is the perimeter of $\mathcal{C}$. On the other hand, since $\partial \mathcal{C}$ is smooth and $\|\mathbf{n}_{\partial \mathcal{C}}(x)\| = 1$ for every x, $\mathbf{n}_{\partial \mathcal{C}}$ can be extended to feasible vector field on Ω (i.e. one that is in $D(\Omega, \mathbb{R}^n)$). Therefore, we get that

$$\mathrm{TV}(u) = \int_{\partial \mathcal{C}} \langle \varphi(x), \mathbf{n}_{\partial \mathcal{C}}(x) \rangle \; dl \geqslant \int_{\partial \mathcal{C}} \|\mathbf{n}_{\partial \mathcal{C}}(x)\|^2 \, dl = \int_{\partial \mathcal{C}} 1 \cdot \; dl = \mathrm{Per}(\mathcal{C}),$$

Therefore, $\mathrm{TV}(\mathbf{1}_{\mathcal{C}}) = \mathrm{Per}_{\mathcal{C}}$ for any domain with smooth boundary. This can be extended to domains with Lipschitz boundary by constructing a sequence of functions in $D(\Omega, \mathbb{R}^n)$ that converge pointwise to $\mathbf{n}_{\partial \mathcal{C}}$.

We now study properties of functions that have a finite value of TV.

Definition 4.3.7. *The functions $u \in L^1(\Omega)$ with a finite value of* TV *form a normed space called the space of functions of bounded variation (the* BV*-space) defined as follows*

$$\mathrm{BV}(\Omega) := \Big\{ u \in L^1(\Omega) \;\Big|\; \|u\|_{\mathrm{BV}} := \|u\|_{L^1} + \mathrm{TV}(u) < \infty \Big\}.$$

Remark 4.3.8. It can be shown that the space BV is the dual of a separable Banach space [13] and that weak-* convergence $u_n \rightharpoonup^* u$ in BV is equivalent to strong convergence $u_n \to u$ in L^1 and convergence of the values $\mathrm{TV}(u_n) \to \mathrm{TV}(u)$. The proof is outside the scope of these notes.

We note that $\mathrm{BV}(\Omega)$ is compactly embedded in $L^1(\Omega)$. We start with the following classical result.

Theorem 4.3.9 (Rellich-Kondrachov, [2, Thm. 6.3])**.** *Let $\Omega \subset \mathbb{R}^n$ be a bounded Lipschitz domain (i.e. non-empty, open, connected and with Lipschitz boundary) and $p, m \in \mathbb{N}$. Let*

$$p^* := \begin{cases} \frac{np}{n-mp} & \text{if } n > mp, \\ \infty & \text{if } n \leqslant mp. \end{cases}$$

Then the embedding $W^{m,p}(\Omega) \to L^q(\Omega)$ is continuous for all $1 \leqslant q \leqslant p^*$ and compact for all $1 \leqslant q < p^*$.

Since functions from $\mathrm{BV}(\Omega)$ can be approximated by functions in the Sobolev space $W^{1,1}(\Omega)$ [5, Thm. 3.9], the Rellich-Kondrachov Theorem (with $p = 1$, $m = 1$) gives us the following

Corollary 4.3.10 ([5, Corrollary 3.49])**.** For any bounded Lipschitz domain $\Omega \subset \mathbb{R}^n$, the embedding

$$\mathrm{BV}(\Omega) \subset\subset L^1(\Omega)$$

is compact for any $n \geqslant 2$ and the embedding

$$\mathrm{BV}(\Omega) \hookrightarrow L^2(\Omega)$$

is continuous for $n = 2$.

Now we will show that TV is lower-semicontinuous in L^1.

Theorem 4.3.11. *Let $\Omega \subset \mathbb{R}^n$ be open and bounded. Then the total variation is l.s.c. in $L^1(\Omega)$.*

Proof. Let $\{u_j\}_{j\in\mathbb{N}} \subset \mathrm{BV}(\Omega)$ be a sequence converging in $L^1(\Omega)$ with $u_j \to u$ in $L^1(\Omega)$. Then for any test function $\varphi \in \mathcal{D}(\Omega, \mathbb{R}^n)$ we have that

$$\int_\Omega u_j(x) \operatorname{div} \varphi(x) dx \to \int_\Omega u(x) \operatorname{div} \varphi(x) dx$$

(strong convergence implies weak convergence) and therefore

$$\begin{aligned}
\mathrm{TV}(u) &= \sup_{\varphi\in\mathcal{D}(\Omega,\mathbb{R}^n)} \int_\Omega u(x) \operatorname{div} \varphi(x) dx \\
&= \sup_{\varphi\in\mathcal{D}(\Omega,\mathbb{R}^n)} \lim_{j\to\infty} \int_\Omega u_j(x) \operatorname{div} \varphi(x) dx \\
&\leqslant \liminf_{j\to\infty} \sup_{\varphi\in\mathcal{D}(\Omega,\mathbb{R}^n)} \int_\Omega u_j(x) \operatorname{div} \varphi(x) dx \\
&= \liminf_{j\to\infty} \mathrm{TV}(u_j).
\end{aligned}$$

Here the lim inf appears when we swap the sup and the lim, because the limit of the suprema may not exist; however, the inequality holds for any subsequence and hence also for the lim inf. Note also that the left and right hand sides may not be finite. □

Since the null space of total variation (4.12) is nontrivial, TV cannot be coercive on L^1. However, the following result helps.

Proposition 4.3.12 ([5, Remark 3.50])**.** *Let $\Omega \subset \mathbb{R}^n$ be a bounded Lipschitz domain. Then there exists a constant $C > 0$ such that for all $u \in \mathrm{BV}(\Omega)$ the* Poincaré *inequality is satisfied*

$$\|u - u_\Omega\|_{L^1} \leqslant C\, \mathrm{TV}(u),$$

where $u_\Omega := \frac{1}{|\Omega|} \int_\Omega u(x)dx$ is the mean-value of u over Ω.

Corollary 4.3.13. It is often useful to consider a subspace $\mathrm{BV}_0(\Omega) \subset \mathrm{BV}(\Omega)$ of functions with zero mean, i.e.

$$\mathrm{BV}_0(\Omega) := \{u \in \mathrm{BV}(\Omega) \colon \int_\Omega u(x)dx = 0\}. \tag{4.13}$$

Then for every function $u \in \mathrm{BV}_0(\Omega)$ we have that

$$\|u\|_{L^1} \leqslant C\, \mathrm{TV}(u).$$

Clearly, $\mathrm{BV}_0 \subset L^1_0 := \{u \in L^1 \colon \int_\omega u(x)\, dx = 0\}$ in TV is coercive on this subspace. Since $\dim(\mathcal{N}(\mathrm{TV})) = 1 < \infty$, we have

$$L^1 = L^1_0 \oplus \mathcal{N}(\mathrm{TV}).$$

Combining all the above results we get

Theorem 4.3.14. *Let $\mathcal{X} = L^1(\Omega)$, where $\Omega \subset \mathbb{R}^n$ is bounded Lipschitz, and $\mathcal{Y}$ be a Banach space. Let $A\colon L^1 \to \mathcal{Y}$ be a linear bounded operator such that $A\mathbf{1} \neq 0$, where $\mathbf{1}$ is the constant-one function. Then minimisers of the following problem*

$$\min_{u \in L^1(\Omega)} \frac{1}{2}\|Au - f_\delta\|_{\mathcal{Y}}^2 + \alpha(\delta)\,\mathrm{TV}(u)$$

converge strongly in L^1 to a TV*-minimising solution as $\delta \to 0$ if $\alpha(\delta)$ is chosen as required by Theorem 4.2.7.*

Proof. We have established all ingredients required for Theorem 4.2.7 to hold except that bounded sequences in L^1 may not have convergent subsequences (L^1 is not a dual space). However, the compact embedding from Corollary 4.3.10 guarantees that sequences with a bounded value of TV have subsequences that converge strongly in L^1. □

Remark 4.3.15. One can replace optimisation over $u \in L^1$ with optimisation over $u \in \mathrm{BV}$, which is the effective domain of the objective function.

Total Variation is widely used in imaging applications [34]. The so-called Rudin–Osher–Fatemi (ROF) model for image denoising [31] consists in minimising the following functional

$$\min_{u \in \mathrm{BV}(\Omega)} \frac{1}{2}\|Iu - f_\delta\|_{L^2(\Omega)}^2 + \alpha\,\mathrm{TV}(u), \tag{4.14}$$

where $\Omega \subset \mathbb{R}^2$. In this case, the forward operator I is the embedding operator $\mathrm{BV}(\Omega) \to L^2(\Omega)$, which is continuous for two-dimensional domains (see Corollary 4.3.10). Clearly, $A\mathbf{1} \neq 0$ is satisfied. More generally, one considers the following optimisation problem

$$\min_{u \in \mathrm{BV}(\Omega)} \|Au - f_\delta\|_2^2 + \alpha\,\mathrm{TV}(u), \tag{4.15}$$

where $A\colon \mathrm{BV}(\Omega) \to L^2(\Omega)$ is such that $A\mathbf{1} \neq 0$.

Chapter 5

Convex Duality

In Chapter 4 we have established convergence of a regularised solution u_δ to a $\mathcal{J}$-minimising solution $u^\dagger_{\mathcal{J}}$ as $\delta \to 0$. However, we didn't get any results on the *speed* of this convergence, which is referred to as the *convergence rate*.

In modern regularisation methods, convergence rates are usually studied using *Bregman distances* associated with the (convex) regularisation functional $\mathcal{J}$. Recall that for a convex functional $\mathcal{J}$, $u, v \in \mathcal{X}$ such that $\mathcal{J}(v) < \infty$ and $q \in \partial\mathcal{J}(v)$, the (generalised) Bregman distance is given by the following expression (cf. Def. 4.1.24)

$$D^q_{\mathcal{J}}(u, v) = \mathcal{J}(u) - \mathcal{J}(v) - \langle q, u - v\rangle \ .$$

Also widely used is the *symmetric* Bregman distance (cf. Def. 4.1.26) given by the following expression (here $p \in \partial\mathcal{J}(u)$)

$$D^{symm}_{\mathcal{J}}(u, v) = D^q_{\mathcal{J}}(u, v) + D^p_{\mathcal{J}}(v, u) = \langle p - q, u - v\rangle \ .$$

Bregman distances appear to be a natural distance measure between a regularised solution u_δ and a $\mathcal{J}$-minimising solution $u^\dagger_{\mathcal{J}}$. For instance, for classical Hilbert space regularisation with $\mathcal{J}(u) = \frac{1}{2}\|u\|^2_{\mathcal{X}}$, the subgradient at $u^\dagger_{\mathcal{J}}$ is $p_{u^\dagger_{\mathcal{J}}} = u^\dagger_{\mathcal{J}}$ (since $\mathcal{J}$ is differentiable) and we get the following expression

$$\begin{aligned} D^{u^\dagger_{\mathcal{J}}}_{\mathcal{J}}(u_\delta, u^\dagger_{\mathcal{J}}) &= \frac{1}{2}\|u_\delta\|^2_{\mathcal{X}} - \frac{1}{2}\|u^\dagger_{\mathcal{J}}\|^2_{\mathcal{X}} - \left\langle u^\dagger_{\mathcal{J}}, u_\delta - u^\dagger_{\mathcal{J}}\right\rangle \\ &= \frac{1}{2}(\|u_\delta\|^2_{\mathcal{X}} - 2\left\langle u^\dagger_{\mathcal{J}}, u_\delta\right\rangle + \|u^\dagger_{\mathcal{J}}\|^2_{\mathcal{X}}) = \frac{1}{2}\|u_\delta - u^\dagger_{\mathcal{J}}\|^2_{\mathcal{X}}, \end{aligned}$$

which happens to coincide with the symmetric Bregman distance. Therefore, in the classical L^2-case, the Bregman distance just measures the L^2-distance between a regularised solution and a $\mathcal{J}$-minimising solution. As we have seen in an example sheet, subgradients of absolutely one-homogeneous functional carry structural information about the solution such as locations of non-zero components of a vector $u^\dagger_{\mathcal{J}} \in \ell^1$.

We are looking for a convergence rate of the following form

$$D^{symm}_{\mathcal{J}}(u_\delta, u^\dagger_{\mathcal{J}}) \leqslant \psi(\delta),$$

where $\psi\colon \mathbb{R}_+ \to \mathbb{R}_+$ is a known function of δ such that $\psi(\delta) \to 0$ as $\delta \to 0$.

5.1 Dual Problem

Recall that u_δ solves the following problem

$$\min_{u\in\mathcal{X}} \frac{1}{2}\|Au - f_\delta\|_{\mathcal{Y}}^2 + \alpha\mathcal{J}(u). \tag{5.1}$$

with an appropriately chosen $\alpha = \alpha(\delta)$, where $\mathcal{X}$ and $\mathcal{Y}$ are Banach spaces, $A \in \mathcal{L}(\mathcal{X},\mathcal{Y})$ and $E\colon \mathcal{Y} \to \bar{\mathbb{R}}$ and $\mathcal{J}\colon \mathcal{X} \to \bar{\mathbb{R}}$ is proper, convex and l.s.c. and satisfies Assumptions of Theorem 4.2.6. For simplicity of presentation, we will also assume throughout this chapter that $\mathcal{J}$ is absolutely one-homogeneous and that $\inf\{\|Av - f\|\colon v \in \mathcal{X}\} = 0$, i.e. $Au_{\mathcal{J}}^\dagger = f$ for any $\mathcal{J}$-minimising solution.

To apply the results of Section 4.1.4 to (5.1), we take (in the notation of Section 4.1.4)

$$E(y) := \frac{1}{2}\|y - f\|_{\mathcal{Y}}^2, \quad F(u) := \alpha\mathcal{J}(u).$$

Lemma 5.1.1. *Let X be a Banach space with norm $\|\cdot\|_X$ and let $\|\cdot\|_{X*}$ be the norm in the dual space of X. Let $\varphi(x) := \frac{1}{2}\|x\|_X^2$. Then the convex conjugate of φ is*

$$\varphi^*(\xi) = \frac{1}{2}\|\xi\|_{X^*}^2, \quad \xi \in X^*.$$

Proof. First, we note that

$$\varphi^*(\xi) = \sup_{x\in X} \langle \xi, x\rangle - \frac{1}{2}\|x\|_X^2 \leqslant \sup_{x\in X} \|x\|_X\|\xi\|_{X^*} - \frac{1}{2}\|x\|_X^2.$$

The function on the right-hand side is a parabola in the scalar variable $\|x\|_X$ and its maximum is $\frac{1}{2}\|\xi\|_{X^*}^2$. Now, fix $\xi \in X^*$. We have that

$$\|\xi\|_{X^*} = \sup_{\substack{x\in X\\ \|x\|=1}} \langle \xi, x\rangle = \sup_{\substack{x\in X\\ \|x\|=\|\xi\|}} \frac{\langle \xi, x\rangle}{\|\xi\|}.$$

Let $x_n^\xi \in X$ be a maximising sequence (that is, $\|x_n^\xi\| = \|\xi\|$ and $\langle \xi, x_n^\xi\rangle \to \|\xi\|^2$). Then

$$\varphi^*(\xi) = \sup_{x\in X} \langle \xi, x\rangle - \frac{1}{2}\|x\|_X^2 \geqslant \limsup_{n\to\infty} \left(\langle \xi, x_n^\xi\rangle - \frac{1}{2}\|x_n^\xi\|_X^2 \right) = \|\xi\|^2 - \frac{1}{2}\|\xi\|^2 = \frac{1}{2}\|\xi\|^2.$$

The inequality here is due to the fact that the lim sup is a supremum over a smaller set than the whole X. Hence, we have that $\frac{1}{2}\|\xi\|^2 \leqslant \varphi^*(\xi) \leqslant \frac{1}{2}\|\xi\|^2$ and the proof is complete. □

Corollary 5.1.2. Theorem 4.1.23 implies that for any $x \in X$ and any $\xi \in \partial\varphi(x)$ it holds

$$\frac{1}{2}\|x\|_X^2 + \frac{1}{2}\|\xi\|_{X^*}^2 = \langle \xi, x\rangle.$$

Using the Cauchy-Schwarz inequality on the right-hand side and rearranging terms, we get that $(\|x\|_X - \|\xi\|_{X^*})^2 = 0$ and hence

$$\|\xi\|_{X^*} = \|x\|_X.$$

Now, for E and F as defined above, we get

$$\begin{aligned} E^*(\eta) &= \sup_{y\in\mathcal{Y}} \langle \eta, y\rangle - \frac{1}{2}\|y-f\|_{\mathcal{Y}}^2 = \langle \eta, f\rangle - \sup_{z\in\mathcal{Y}} \left(\langle \eta, z\rangle - \frac{1}{2}\|z\|_{\mathcal{Y}}^2 \right) = \langle \eta, f\rangle + \frac{1}{2}\|\eta\|_{\mathcal{Y}^*}, \\ F^*(p) &= \chi_{\partial \mathcal{J}(0)} \left(\frac{p}{\alpha} \right), \end{aligned}$$

where the second equality holds since F is absolutely one-homogeneous. Hence, the dual problem of (5.1) is given by

$$\sup_{\eta\in\mathcal{Y}^*} -\langle \eta, f\rangle - \frac{1}{2}\|\eta\|_{\mathcal{Y}^*}^2 - \chi_{\partial\mathcal{J}(0)} \left(\frac{-A^*\eta}{\alpha} \right).$$

Denote $\mu := -\frac{\eta}{\alpha} \in \mathcal{Y}^*$. Since $-\chi_{\partial\mathcal{J}(0)} = -\infty$ outside $\partial\mathcal{J}(0)$, we get the following equivalent problem

$$\sup_{\substack{\mu\in\mathcal{Y}^* \\ A^*\mu\in\partial\mathcal{J}(0)}} \alpha \left(\langle \mu, f\rangle - \frac{\alpha}{2}\|\mu\|_{\mathcal{Y}}^2 \right). \tag{5.2}$$

Let us check if Assumptions of Theorem 4.1.43 are satisfied. Condition (i) (coercivity) is guaranteed by Lemma 4.2.5. Condition (ii) (continuity of E) is satisfied at $u_0 = 0$. Therefore, for any $\delta > 0$ there exists a solution μ_δ of the dual problem (5.2).

Existence of a primal solution u_δ is guaranteed by Theorem 4.2.6. Indeed, let us take $\tau_{\mathcal{X}}$ to be the weak* topology in $\mathcal{X}$ and $\tau_{\mathcal{Y}}$ a topology in $\mathcal{Y}$ such that A is $\tau_{\mathcal{X}}$-$\tau_{\mathcal{Y}}$ continuous and the norm in $\mathcal{Y}$ is $\tau_{\mathcal{Y}}$-l.s.c. (weak*, weak or strong topologies will work). For example, if $\mathcal{Y}$ has a separable predual, we can take $\tau_{\mathcal{Y}}$ to be the weak* topology on $\mathcal{Y}$. It can be easily verified that A is weak*-weak* continuous if it is the dual of another operator $A = B^*$ (where B acts from the predual of $\mathcal{Y}$ into the predual of $\mathcal{X}$). With these choices, the conditions of Theorem 4.2.6 are satisfied.

Hence, by strong duality we have that

$$\frac{1}{2}\|Au_\delta - f_\delta\|_{\mathcal{Y}}^2 + \alpha\mathcal{J}(u_\delta) = \alpha \langle \mu_\delta, f_\delta\rangle - \frac{\alpha^2}{2}\|\mu_\delta\|_{\mathcal{Y}}^2.$$

Optimality conditions (iii) from Theorem 4.1.43 take the following form

$$A^*\mu_\delta \in \partial\mathcal{J}(u_\delta), \quad -\alpha\mu_\delta \in \partial \left(\frac{1}{2}\|\cdot\|_{\mathcal{Y}}^2 \right)(Au_\delta - f_\delta). \tag{5.3}$$

From Corollary 5.1.2 we conclude that

$$\|\alpha\mu_\delta\|_{\mathcal{Y}^*} = \|Au_\delta - f_\delta\|_{\mathcal{Y}}. \tag{5.4}$$

Also, comparing the values of $\frac{1}{2}\|\cdot\|^2$ at 0 and at $Au_\delta - f_\delta$ and using the fact that $-\alpha\mu_\delta$ is a subgradient, we get that

$$0 \geqslant \frac{1}{2}\|Au_\delta - f_\delta\|_{\mathcal{Y}}^2 + \langle -\alpha\mu_\delta, 0 - (Au_\delta - f_\delta)\rangle$$

and therefore

$$\langle \alpha\mu_\delta, Au_\delta - f_\delta\rangle \leqslant -\frac{1}{2}\|Au_\delta - f_\delta\|_{\mathcal{Y}}^2. \tag{5.5}$$

We will use the estimates (5.4) and (5.5) later in Theorem 5.2.4.

5.2 Source Condition and Convergence Rates

Formal limits of problems (5.1) and (5.2) at $\delta = 0$ are

$$\inf_{u\colon Au=f} \mathcal{J}(u) = \inf_{u\in\mathcal{X}} \chi_{\{f\}}(Au) + \mathcal{J}(u) \tag{5.6}$$

and

$$\begin{aligned}\sup_{\mu\colon A^*\mu\in\partial\mathcal{J}(0)} \langle \mu, f\rangle = & \sup_{\mu\colon A^*\mu\in\partial\mathcal{J}(0)} \left\langle \mu, Au^\dagger_{\mathcal{J}}\right\rangle \\ = & \sup_{\mu\colon A^*\mu\in\partial\mathcal{J}(0)} \left\langle A^*\mu, u^\dagger_{\mathcal{J}}\right\rangle = \sup_{v\in\mathcal{R}(A^*)\cap\partial\mathcal{J}(0)} \left\langle v, u^\dagger_{\mathcal{J}}\right\rangle. \end{aligned} \tag{5.7}$$

Since the characteristic function $\chi_{\{f\}}(\cdot)$ is not continuous anywhere in its domain, Theorem 4.1.43 does not apply and we cannot guarantee that a solution of the dual limit problem (5.7) exists. Indeed, since $\mathcal{R}(A^*)$ is not closed (strongly and hence weakly, since it is convex [18, Thm. V.3.13]), a solution may not exist.

We shall see that existence is guaranteed by the following condition

Definition 5.2.1 (Source condition [14])**.** *We say that a $\mathcal{J}$-minimising solution $u^\dagger_{\mathcal{J}}$ satisfies the* source condition *if*

$$\exists \mu^\dagger \in \mathcal{Y}^* \quad \textit{such that} \quad A^*\mu^\dagger \in \partial\mathcal{J}(u^\dagger_{\mathcal{J}}), \tag{5.8}$$

i.e. if $\mathcal{R}(A^*) \cap \partial\mathcal{J}(u^\dagger_{\mathcal{J}}) \neq \emptyset$.

First we will see that this condition is necessary for the dual solution μ_δ from (5.3) to stay bounded as $\delta \to 0$.

Theorem 5.2.2 (Necessary conditions, [24])**.** *Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces and $\mathcal{Y}$ separable. Let conditions of Theorem 4.2.6 be satisfied and $\alpha = \alpha(\delta)$ be chosen as required by Theorem 4.2.7. Suppose that the dual solution μ_δ is bounded uniformly in δ. Then there exists $\mu^\dagger \in \mathcal{Y}^*$ such that $A^*\mu^\dagger \in \partial\mathcal{J}(u^\dagger_{\mathcal{J}})$.*

Proof. Consider an arbitrary sequence $\delta_n \downarrow 0$. Since $\|\mu_\delta\|_{\mathcal{Y}^*} \leqslant C$ for all δ, by the Banach-Alaogly theorem we get that there exists a weakly-* convergent subsequence (that we do not relabel), i.e.

$$\mu_{\delta_n} \rightharpoonup^* \mu^\dagger \in \mathcal{Y}^*.$$

Then we get that

$$A^*\mu_{\delta_n} \rightharpoonup^* A^*\mu^\dagger.$$

Since $\partial\mathcal{J}(0)$ is weakly-* closed (Theorem 4.1.19) and $A^*\mu_{\delta_n} \in \partial\mathcal{J}(0)$ by (5.3), we get that

$$A^*\mu^\dagger \in \partial\mathcal{J}(0).$$

Since $\mathcal{J}$ is absolute one-homogeneous, we get by Proposition 4.1.28 that

$$\langle A^*\mu_{\delta_n}, u_{\delta_n}\rangle = \mathcal{J}(u_{\delta_n}) \to \mathcal{J}(u^\dagger_{\mathcal{J}}), \tag{5.9}$$

where convergence follows from Theorem 4.2.7. We also observe that

$$\begin{aligned} |\langle A^*\mu_\delta, u_\delta\rangle - \langle A^*\mu^\dagger, u^\dagger_{\mathcal{J}}\rangle| &= |\langle A^*\mu_\delta, u_\delta - u^\dagger_{\mathcal{J}}\rangle - \langle A^*(\mu^\dagger - \mu_\delta), u^\dagger_{\mathcal{J}}\rangle| \\ &\leqslant |\langle \mu_\delta, Au_\delta - f\rangle| + |\langle \mu^\dagger - \mu_\delta, f\rangle| \\ &\leqslant \|\mu_\delta\|\|Au_\delta - f\| + |\langle \mu^\dagger - \mu_\delta, f\rangle| \to 0, \end{aligned}$$

since $\|\mu_{\delta_n}\|_{\mathcal{Y}^*}$ is bounded, $\|Au_{\delta_n} - f\|_{\mathcal{Y}} \to 0$ and $\mu_{\delta_n} \rightharpoonup^* \mu^\dagger$. Combining this with (5.9), we get that

$$\mathcal{J}(u^\dagger_{\mathcal{J}}) = \left\langle A^*\mu^\dagger, u^\dagger_{\mathcal{J}} \right\rangle.$$

Since $A^*\mu^\dagger \in \partial\mathcal{J}(0)$ and $\mathcal{J}(u^\dagger_{\mathcal{J}}) = \left\langle A^*\mu^\dagger, u^\dagger_{\mathcal{J}} \right\rangle$, we conclude, using Proposition 4.1.31, that $A^*\mu^\dagger \in \partial\mathcal{J}(u^\dagger_{\mathcal{J}})$. □

So, the source condition is necessary for the boundedness of the dual solutions μ_δ as $\delta \to 0$. It turns out to be also sufficient.

Theorem 5.2.3 (Sufficient conditions, [24])**.** *Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces and $\mathcal{Y}$ separable. Let conditions of Theorem 4.2.6 be satisfied and $\alpha = \alpha(\delta)$ be chosen as required by Theorem 4.2.7. Suppose that the source condition* (5.8) *is satisfied at a $\mathcal{J}$-minimising solution $u^\dagger_{\mathcal{J}}$. Then μ_δ is bounded uniformly in δ. Moreover, $\mu_\delta \rightharpoonup^* \mu^\dagger$ in $\mathcal{Y}^*$ as $\delta \to 0$ (perhaps, up to a subsequence), where $\mu^\dagger$ is the solution of the dual limit problem* (5.7) *with minimal norm.*

Proof. We omit the proof for time reasons. It can be found in [24] (for Hilbert spaces). □

The next theorem shows that the source condition (5.8) implies a convergence rates in terms of the Bregman distance.

Theorem 5.2.4. *Let the source condition* (5.8) *be satisfied at a $\mathcal{J}$-minimising solution $u^\dagger_{\mathcal{J}}$ and let u_δ be a regularised solution solving* (5.1)*. Then the following estimate holds*

$$D^{p_\delta,p^\dagger}_{\mathcal{J}}(u_\delta, u^\dagger_{\mathcal{J}}) \leqslant \frac{1}{4\alpha}\left(\delta + \alpha\|\mu^\dagger\|\right)^2 + \delta\|\mu^\dagger\|.$$

where $p_\delta = A^*\mu_\delta \in \partial\mathcal{J}(u_\delta)$ with μ_δ as defined in (5.3) *and $p^\dagger = A^*\mu^\dagger \in \partial\mathcal{J}(u^\dagger_{\mathcal{J}})$ is as defined in* (5.8)*. $D^{p_\delta,p^\dagger}_{\mathcal{J}}(u_\delta, u^\dagger_{\mathcal{J}})$ denotes the symmetric Bregman distance between u_δ and $u^\dagger_{\mathcal{J}}$. For the optimal choice $\alpha = \frac{\delta}{\|\mu^\dagger\|}$ we get that*

$$D^{p_\delta,p^\dagger}_{\mathcal{J}}(u_\delta, u^\dagger_{\mathcal{J}}) \leqslant 3\delta\|\mu^\dagger\|.$$

Proof. We start with the following estimate

$$\begin{aligned}
\alpha D^{p_\delta,p^\dagger}_{\mathcal{J}}(u_\delta, u^\dagger_{\mathcal{J}}) &= \alpha\langle p_\delta - p^\dagger, u_\delta - u^\dagger_{\mathcal{J}}\rangle \\
&= \alpha\langle \mu_\delta - \mu^\dagger, Au_\delta - f\rangle \\
&= \alpha\langle \mu_\delta, Au_\delta - f_\delta\rangle + \alpha\langle \mu_\delta, f_\delta - f\rangle - \alpha\langle \mu^\dagger, Au_\delta - f_\delta\rangle - \alpha\langle \mu^\dagger, f_\delta - f\rangle.
\end{aligned}$$

From (5.5) we know that

$$\alpha\langle \mu_\delta, Au_\delta - f_\delta\rangle \leqslant -\frac{1}{2}\|Au_\delta - f_\delta\|^2_{\mathcal{Y}}.$$

and from (5.4) that $\alpha\|\mu_\delta\| = \|Au_\delta - f_\delta\|$. Using these estimates, the Cauchy-Schwarz inequality and the estimate $\|f - f_\delta\| \leqslant \delta$, we get

$$\alpha D^{p_\delta,p^\dagger}_{\mathcal{J}}(u_\delta, u^\dagger_{\mathcal{J}}) \leqslant -\frac{1}{2}\|Au_\delta - f_\delta\|^2 + \left(\delta + \alpha\|\mu^\dagger\|\right)\|Au_\delta - f_\delta\| + \alpha\delta\|\mu^\dagger\|.$$

The right-hand side is the following quadratic function of the scalar variable $\|Au_\delta - f_\delta\|$

$$\varphi(t) := -\frac{1}{2}t^2 + (\delta + \alpha\|\mu^\dagger\|)t + \alpha\delta\|\mu^\dagger\|, \quad t \in \mathbb{R}.$$

It achieves its maximum at $t_0 = (\delta + \alpha\|\mu^\dagger\|)$ and this maximum value is equal to

$$\varphi(t_0) = \frac{(\delta + \alpha\|\mu^\dagger\|)^2}{2} + \alpha\delta\|\mu^\dagger\|.$$

Substituting this into the above estimate for the Bregman distance and dividing both sides by α, we get the desired estimate

$$D_{\mathcal{J}}^{p_\delta, p^\dagger}(u_\delta, u_{\mathcal{J}}^\dagger) \leqslant \frac{(\delta + \alpha\|\mu^\dagger\|)^2}{2\alpha} + \delta\|\mu^\dagger\|.$$

Differentiating the right-hand side w.r.t. α and setting the derivative to zero, we obtain the following optimality condition for α

$$0 = \frac{2\alpha\|\mu^\dagger\|(\delta + \alpha\|\mu^\dagger\|) - (\delta + \alpha\|\mu^\dagger\|)^2}{2\alpha^2} = \frac{\alpha^2\|\mu^\dagger\|^2 - \delta^2}{2\alpha^2}$$

and

$$\alpha = \frac{\delta}{\|\mu^\dagger\|}.$$

With this optimal choice of α we get the following estimate

$$D_{\mathcal{J}}^{p_\delta, p^\dagger}(u_\delta, u_{\mathcal{J}}^\dagger) \leqslant 3\delta\|\mu^\dagger\|.$$

□

Remark 5.2.5. Of course, we do not know $\mu^\dagger$ since we don't know the $\mathcal{J}$-minimising solution $u_{\mathcal{J}}^\dagger$, but the theorem gives an optimal *rate* $\alpha \sim \delta$ for a priori parameter choice rules and a corresponding error estimate $D_{\mathcal{J}}^{p_\delta, p^\dagger}(u_\delta, u_{\mathcal{J}}^\dagger) = O(\delta)$.

Now we will look at two examples involving Total Variation to get a feeling for what the source condition 'means'.

Example 5.2.6 (Total Variation)**.** Let $\Omega \subset \mathbb{R}^2$ be a bounded domain with a C^∞ boundary. Let $\mathcal{X} = \mathrm{BV}(\Omega)$ and $\mathcal{Y} = L^2(\Omega)$ and $\mathcal{J}(\cdot) = \mathrm{TV}(\cdot)$. Recall the ROF problem

$$\min_{u \in \mathrm{BV}} \frac{1}{2}\|Iu - f_\delta\|_{L^2}^2 + \alpha\,\mathrm{TV}(u),$$

where $I\colon \mathrm{BV}(\Omega) \to L^2(\Omega)$ is the embedding operator, which is continuous since $\Omega \subset \mathbb{R}^2$. The adjoint $I^*\colon L^2(\Omega) \to \mathrm{BV}^*(\Omega)$ continuously embeds L^2 into BV^*. Clearly, I^* is not surjectuve and $\mathcal{R}(I^*) = L^2(\Omega)$.

From Example 4.3.6 we know that

$$\mathrm{TV}(\mathbf{1}_{\mathcal{C}}) = \mathrm{Per}(\mathcal{C}),$$

where $\mathbf{1}_{\mathcal{C}}$ is the indicator function of the set $\mathcal{C}$. Denoting by $\mathbf{n}_{\partial\mathcal{C}}$ the unit normal, we obtain

$$\mathrm{Per}(\mathcal{C}) = \int_{\partial\mathcal{C}} 1 = \int_{\partial\mathcal{C}} \langle \mathbf{n}_{\partial\mathcal{C}}, \mathbf{n}_{\partial\mathcal{C}} \rangle .$$

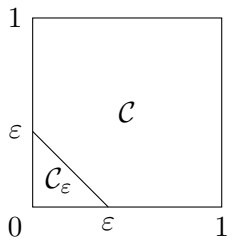


Figure 5.1: Example of a set whose indicator function does not satisfy the source condition.

Since $\mathbf{n}_{\partial\mathcal{C}} \in C^\infty(\partial\mathcal{C}, \mathbb{R}^2)$ and $\|\mathbf{n}_{\partial\mathcal{C}}(x)\|_2 = 1$ for any x, we can extend $\mathbf{n}_{\partial\mathcal{C}}$ to a $C_0^\infty(\Omega, \mathbb{R}^2)$ vector field ψ with $\sup_{x\in\Omega} \|\psi(x)\|_2 \leqslant 1$. Therefore, using the divergence theorem, we obtain that

$$\int_{\partial\mathcal{C}} \langle \mathbf{n}_{\partial\mathcal{C}}, \mathbf{n}_{\partial\mathcal{C}} \rangle = \int_{\partial\mathcal{C}} \langle \psi, \mathbf{n}_{\partial\mathcal{C}} \rangle = \int_{\mathcal{C}} \operatorname{div} \psi = \int_\Omega \mathbf{1}_\mathcal{C} \operatorname{div} \psi.$$

Combining all these equalities, we get that

$$\operatorname{TV}(\mathbf{1}_\mathcal{C}) = \int_\Omega \mathbf{1}_\mathcal{C} \operatorname{div} \psi = \langle \operatorname{div} \psi, \mathbf{1}_\mathcal{C} \rangle.$$

Taking an arbitrary $u \in \operatorname{BV}(\Omega)$, we note that

$$\operatorname{TV}(u) - \langle \operatorname{div} \psi, u \rangle = \sup_{\substack{\varphi \in C_0^\infty(\Omega, \mathbb{R}^2) \\ \sup_{x\in\Omega} \|\varphi(x)\|_2 \leqslant 1}} \langle \operatorname{div} \varphi, u \rangle - \langle \operatorname{div} \psi, u \rangle \geqslant 0,$$

since $\varphi = \psi$ is feasible. Therefore, $\operatorname{div} \psi \in \partial \operatorname{TV}(0)$ and, since $\operatorname{TV}(\mathbf{1}_\mathcal{C}) = \langle \operatorname{div} \psi, \mathbf{1}_\mathcal{C} \rangle$, we also get that

$$\operatorname{div} \psi \in \partial \operatorname{TV}(\mathbf{1}_\mathcal{C}).$$

Since $\psi \in C_0^\infty(\Omega, \mathbb{R}^2)$, we have $\operatorname{div} \psi \in C_0^\infty(\Omega) \subset L^2(\Omega) = \mathcal{R}(I^*)$ and the source condition is satisfied at $u = \mathbf{1}_\mathcal{C}$ with $\mu^\dagger = \operatorname{div} \psi$.

Example 5.2.7 (Total Variation)**.** In the same setting as in Example 5.2.6, let $\mathcal{C}$ be a domain with a nonsmooth boundary, e.g., a square $\mathcal{C} = [0, 1]^2$. We will show in this example that in this case $\partial \operatorname{TV}(\mathbf{1}_\mathcal{C}) \cap \mathcal{R}(I^*) = \emptyset$, where $\mathcal{R}(I^*) = L^2(\Omega)$ as before, i.e. the source condition fails.

Assume that there exists $p_0 \in \partial \operatorname{TV}(\mathbf{1}_\mathcal{C}) \cap L^2(\Omega)$. Then by the results of Example 4.3.6 we have that

$$\langle p_0, \mathbf{1}_\mathcal{C} \rangle = \operatorname{TV}(\mathbf{1}_\mathcal{C}) = \operatorname{Per}(\mathcal{C}) = 4.$$

Since p_0 is a subgradient, we get that for any $u \in \operatorname{BV}(\Omega)$

$$\operatorname{TV}(u) - \langle p_0, u \rangle \geqslant 0.$$

Let us cut a triangle $\mathcal{C}_\varepsilon$ of size ε from a corner of $\mathcal{C}$ as shown in Figure 5.1. Then for $u = \mathbf{1}_{\mathcal{C}\setminus\mathcal{C}_\varepsilon}$ we get

$$\operatorname{TV}(\mathbf{1}_{\mathcal{C}\setminus\mathcal{C}_\varepsilon}) \geqslant \langle p_0, \mathbf{1}_{\mathcal{C}\setminus\mathcal{C}_\varepsilon} \rangle = \langle p_0, \mathbf{1}_\mathcal{C} \rangle - \langle p_0, \mathbf{1}_{\mathcal{C}_\varepsilon} \rangle$$

and therefore

$$\langle p_0, \mathbf{1}_{\mathcal{C}_\varepsilon} \rangle \geqslant \operatorname{TV}(\mathbf{1}_\mathcal{C}) - \operatorname{TV}(\mathbf{1}_{\mathcal{C}\setminus\mathcal{C}_\varepsilon}) = \operatorname{Per}(\mathcal{C}) - \operatorname{Per}(\mathcal{C}\setminus\mathcal{C}_\varepsilon) = 4 - (4 - 2\varepsilon + \sqrt{2}\varepsilon) = (2 - \sqrt{2})\varepsilon > 0.$$

By Hölder's inequality we get that

$$\langle p_0, \mathbf{1}_{\mathcal{C}_\varepsilon} \rangle = \int_{\mathcal{C}_\varepsilon} p_0 \cdot \mathbf{1} \leqslant \left(\int_{\mathcal{C}_\varepsilon} |p_0|^2 \right)^{1/2} \left(\int_{\mathcal{C}_\varepsilon} 1 \right)^{1/2} = \frac{1}{\sqrt{2}} \varepsilon \left(\int_{\mathcal{C}_\varepsilon} |p_0|^2 \right)^{1/2}.$$

Combining the last two inequalities, we get

$$(2 - \sqrt{2})\varepsilon \leqslant \langle p_0, \mathbf{1}_{\mathcal{C}_\varepsilon} \rangle \leqslant \frac{1}{\sqrt{2}} \varepsilon \left(\int_{\mathcal{C}_\varepsilon} |p_0|^2 \right)^{1/2}$$

and therefore

$$\int_{\mathcal{C}_\varepsilon} |p_0|^2 \geqslant 2(2 - \sqrt{2})^2 > 0$$

for all $\varepsilon > 0$. However, since $p_0 \in L^2(\Omega)$ by assumption, we must have

$$\int_{\mathcal{C}_\varepsilon} |p_0|^2 \to 0 \quad \text{as } \varepsilon \to 0.$$

This contradiction proves that such p_0 does not exist and $\partial \operatorname{TV}(\mathbf{1}_{\mathcal{C}}) \cap \mathcal{R}(I^*) = \emptyset$.

Chapter 6

Bayesian probability and statistics

6.1 From inverse problems to Bayesian inverse problems

We consider an inverse problem of the form:

$$\text{Find } u \in \mathcal{X} : \mathcal{A}(u) + n = f_n,$$

where $\mathcal{X}$ is a separable Banach space, $n \in \mathcal{Y}$ is observational noise, $\mathcal{Y}$ is another separable Banach space, $f_n \in \mathcal{Y}$ is data, and $\mathcal{A} : \mathcal{X} \to \mathcal{Y}$ is a measurable (possibly non-linear) operator.

So far, we have studied techniques (pseudo-inverse, regularisation) to find estimates for the parameter u. In situation where the noise n is large or the data is non-informative, we should not only give an estimate for u, but also comment on the uncertainty left in the parameter. This is the problem we study in this part of the lecture.

There are multiple ways to represent certainty, knowledge, risk, or uncertainty in a parameter, such as $u \in \mathcal{X}$. Common models are Bayesian probability theory, fuzzy set theory, Dempster–Shafer theory, random set theory,...

We follow Bayesian probability theory: model uncertain parameters as random variables.

Intuitions, concepts, questions, and answers:

1. Can we use randomness to model deterministic, uncertain objects?
   - Not with the usual "frequentist" interpretation of probability. Here, the probability of an event is the limit of the relative frequency of the occurrence of the event in infinitely repeated, independent experiments. If the object we study is deterministic, the frequentist approach will only give us probabilities in $\{0, 1\}$.
   - Indeed, with the "Bayesian" interpretation of probability. Here the probability of an event is the amount of money (in £) we would give in a game to win £1 if the event occurs. This 'game' does not require any inherent randomness.
2. Can we represent the learning of information about a parameter?
   - Learning that an event B occurred can be represented via conditional probability. Indeed, this learning process is given by the map $\mathbb{P}(U \in \cdot) \mapsto \mathbb{P}(U \in \cdot | B)$.
   - In practice, we can often compute updates of this form through Bayes' formula.

3. Can we use Bayesian probability to argue about logical statements?
   - Cox's Theorem [17]: Bayesian probability is a sensible extension of Aristotelian logic.
4. Is Bayesian probability theory congruent with our everyday experience?
   - It probably is. See the example below.

   Example 6.1.1. 'Tossing a coin' can be modelled as a Bernoulli experiment

   $$\mathbb{P}(\text{Coin shows Head}) = 0.5 = \mathbb{P}(\text{Coin shows Tail}).$$

   Actually, this is a mechanical process that is completely deterministic. However, it is difficult to predict its outcome. The model is complicated and subject to many uncertain parameters: force, speed, gravity, air flow... Hence, it is easier to model the coin as a random variable.

5. How do we employ these ideas in inverse problems?
   (a) We assume that noise n and parameter u are random variables N and U. The distributions of N and U describe our knowledge concerning noise and parameter before observing the data set. The distribution of U is called prior distribution $\mu_0 := \mathbb{P}(U \in \cdot)$.
   (b) We observe the data set f_n, indeed, we observe the occurrence of the event
   $$\{f_n = \mathcal{A}(U) + N\}.$$
   (c) We employ Bayes' theorem to 'update' the prior by incorporating the observational data
   $$\mu_0 = \mathbb{P}(U \in \cdot) \mapsto \mathbb{P}(U \in \cdot | f_n = \mathcal{A}(U) + N) =: \mu_{\text{post}}.$$
   As μ_{post} now explains our knowledge after seeing the data, we call it posterior distribution.

6.2 Reminder: measure, probability, and integration

During this course, we will make extensive use of measure-theoretic probability theory. Thus, we will briefly remind ourselves of some definitions, examples, and results from measure and probability theory that we will require throughout this lecture. In case the reader would like to get a more thorough reminder, we refer them to [6], [10], [25]. We commence with σ-algebras.

Definition 6.2.1 (σ-algebra)**.** *Let Ω be a non-empty set, let $2^\Omega := \{A : A \subseteq \Omega\}$ be the power set of Ω, and let $\mathcal{F} \subseteq 2^\Omega$ satisfy (i)-(iii):*

(i) $\Omega \in \mathcal{F}$,

(ii) for any $F \in \mathcal{F}$, we have also $F^c := \Omega \backslash F \in \mathcal{F}$, and

(iii) for any countable family $(F_n : n \in \mathbb{N}) \in \mathcal{F}^{\mathbb{N}}$, we have also $\bigcup_{n\in\mathbb{N}} F_n \in \mathcal{F}$.

Then, $\mathcal{F}$ is called σ-algebra on Ω and $(\Omega, \mathcal{F})$ is called measurable space.

There are several ways to construct σ-algebras. They can for instance be induced by systems of sets or functions.

Definition 6.2.2 (Induced σ-algebra)**.** *1. Let Ω be non-empty and $\mathcal{E} \subseteq 2^\Omega$. We define the σ-algebra induced by $\mathcal{E}$ on Ω by*

$$\sigma_\Omega(\mathcal{E}) := \bigcap_{\substack{\mathcal{F}' \supset \mathcal{E} \\ \mathcal{F}' \text{ is } \sigma\text{-algebra on } \Omega}} \mathcal{F}'.$$

2. Let Ω be non-empty, let $(\Omega', \mathcal{F}')$ be a measurable space, and let $g : \Omega \to \Omega'$ be a function. We define the σ-algebra induced by g on Ω by

$$\sigma_\Omega(g) := \{\{g \in F'\} : F' \in \mathcal{F}'\},$$

where

$$\{g \in F'\} := g^{-1}(F') := \{\omega \in \Omega : g(\omega) \in F'\}$$

is the pre-image of F' under g.

Example 6.2.1. Let Ω be a non-empty set.

1. 2^Ω is the largest σ-algebra on Ω. $\{\emptyset, \Omega\}$ is the smallest σ-algebra.
2. Let Ω be a topological space with open sets $O \subseteq 2^\Omega$. The σ-algebra $\sigma_\Omega(O) =: \mathcal{B}\Omega$ is called Borel-σ-algebra on Ω.

A σ-algebra is the natural space to define a (probability) measure on.

Definition 6.2.3 (Measure and probability measure)**.** *Let $(\Omega, \mathcal{F})$ be a measurable space and let $\mu : \mathcal{F} \to [0, \infty]$ be a function, satisfying (i),(ii):*

(i) $\mu(\emptyset) = 0$,

(ii) for any countable family $(F_m : m \in \mathbb{N}) \in \mathcal{F}^\mathbb{N}$ of mutually disjoint sets, i.e. $F_n \cap F_m = \emptyset$ $(n \neq m)$. Then, we have $\mu\left(\bigcup_{m\in\mathbb{N}} F_m\right) = \sum_{m\in\mathbb{N}} \mu(F_m)$.

Then, μ is called measure on $(\Omega, \mathcal{F})$ and $(\Omega, \mathcal{F}, \mu)$ is called measure space. If a measure μ additionally satisfies (iii):

(iii) $\mu(\Omega) = 1$,

the measure μ is called probability measure and $(\Omega, \mathcal{F}, \mu)$ is called probability space. Finally, a measure μ is called σ-finite, if

(iv) there is a countable family $(F_m : m \in \mathbb{N}) \in \mathcal{F}^\mathbb{N}$, with $\bigcup_{m\in\mathbb{N}} F_m = \Omega$ and $\mu(F_m) < \infty$ $(m \in \mathbb{N})$.

Example 6.2.2. Let $(\Omega, \mathcal{F})$ be some measurable space.

- $\# : \mathcal{F} \to [0, \infty]$ defined by

$$\#(F) := \begin{cases} \infty, & \text{if } F \text{ is infinite} \\ |F|, & \text{otherwise.} \end{cases} \qquad (F \in \mathcal{F})$$

is a measure and called counting measure,

- Let $\omega \in \Omega$. Then, $\delta(\cdot - \omega) : \mathcal{F} \to [0, \infty]$ defined by

$$\delta(F - \omega) := \begin{cases} 1, & \text{if } F \ni \omega \\ 0, & \text{otherwise} \end{cases} \qquad (F \in \mathcal{F})$$

  is called Dirac measure concentrated in ω. The Dirac measure is a probability measure.

- Let $k \in \mathbb{N}$, $\Omega := \mathbb{R}^k$, and $\lambda_k : \mathcal{B}\mathbb{R}^k \to [0, \infty]$ be the unique measure that satisfies

$$\lambda_k \left(\prod_{i=1}^{k} [a_i, b_i) \right) = \prod_{i=1}^{k} (b_i - a_i),$$

  if $a_i \leqslant b_i$ $(i = 1, ..., k)$. Then λ_k is called k-dimensional Lebesgue measure.

Exercise 6.2.4. 1. Show that the Dirac and counting measure are measures.

2. Show that Dirac and Lebesgue measure are σ-finite.

3. When is the counting measure σ-finite?

We already learned the concept of using a function to construct a σ-algebra. In the following, we would like to use functions to represent uncertainties ('random variables') and use measures to integrate functions. Here, we require the concept of 'measurability'.

Definition 6.2.5. *Let $(\Omega, \mathcal{F})$ and $(\Omega', \mathcal{F}')$ be two measurable spaces and let $g : \Omega \to \Omega'$ be a function.*

1. *g is called measurable, if $\{g \in F'\} \in \mathcal{F}$, for any $F' \in \mathcal{F}'$. In this case, we sometimes write $g : (\Omega, \mathcal{F}) \to (\Omega', \mathcal{F}')$.*

2. *Let g be measurable and μ be a measure on $(\Omega, \mathcal{F})$. Then, we define the push-forward measure $\mu(g \in \cdot)$. If in addition, μ is a probability measure, g is called random variable and $\mu(g \in \cdot)$ is called (probability) distribution of g.*

This rather abstract definition of measurability does not appear to be very instructive in practice. A useful result is the following proposition

Proposition 6.2.6. *Let Ω be a topological space and $g : \Omega \to \mathbb{R}$ be continuous, i.e. for any open $F' \subseteq \mathbb{R}$, the preimage $\{g \in F'\} \subseteq \Omega$ is open as well. Then, $g : (\Omega, \mathcal{B}\Omega) \to (\mathbb{R}, \mathcal{B}\mathbb{R})$ is measurable.*

Proof. Page 36 in [6]. □

Push-forward measures and probability distributions are well-defined measures and probability measures, respectively.

Proposition 6.2.7. *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space, $(\Omega', \mathcal{F}')$ be a measurable spaces, and let $g : (\Omega, \mathcal{F}) \to (\Omega', \mathcal{F}')$ be a measurable function. Then, the pushforward measure $\mu(g \in \cdot)$ is a measure on $(\Omega', \mathcal{F}')$. Moreover, if μ is a probability measure, then so is $\mu(g \in \cdot)$.*

Proof. Exercise. □

Measurability is the basic concept needed to be able to integrate a function with respect to a measure. We start with simple functions.

Definition 6.2.8. *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. A function $g : \Omega \to \mathbb{R}$ is called simple, if there exists an $m \in \mathbb{N}$ and $(F_i : i = 1, ..., m) \in \mathcal{F}^m$, such that*

$$g = \sum_{i=1}^{m} b_i \mathbf{1}_{F_i},$$

for some $b \in \mathbb{R}^m$. Consider the following two assumptions:

(i) $b \in [0, \infty)^m$ or $b \in (-\infty, 0]^m$,

(ii) for any $i \in \{1, ..., m\}$, with $\mu(F_i) = \infty$, we have $b_i = 0$.

If either (i) or (ii) holds, we define the (Lebesgue) integral of g with respect to μ by

$$\int_\Omega g \mathrm{d}\mu := \int_\Omega g(\omega) \mathrm{d}\mu(\omega) := \int_\Omega g(\omega)\mu(\mathrm{d}\omega) := \sum_{i=1; b_i \neq 0}^{m} b_i \mu(F_i).$$

If the expression on the right-hand side is finite, we call g (Lebesgue) integrable.

Exercise 6.2.9. A simple function $g : \Omega \to \mathbb{R}$ is measurable from $(\Omega, \mathcal{F})$ to $(\mathbb{R}, \mathcal{B}\mathbb{R})$.

To define the integral for more general functions g, we will approximate the function by simple functions. This gives us the following definition for the integral.

Definition 6.2.10 (Lebesgue integral)**.** *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and let $g : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}\mathbb{R})$ be measurable and non-negative. Then, we define the (Lebesgue) integral of g by*

$$\int_\Omega g \mathrm{d}\mu := \sup \left\{ \int_\Omega h(\omega) \mathrm{d}\mu(\omega) : 0 \leqslant h \leqslant g, h \text{ is simple} \right\}$$

If the supremum is finite, we call g (Lebesgue) integrable.

In the following proposition, we discuss the fundamental properties of the Lebesgue integral: linearity, monotonicity, and monotonic convergence.

Proposition 6.2.11. *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and let $g, h, g_1, g_2, \ldots : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}\mathbb{R})$ be measurable, non-negative functions. Then:*

1. *If $g \leqslant h$ pointwise, then $\int_\Omega g \mathrm{d}\mu \leqslant \int_\Omega h \mathrm{d}\mu$.*
2. *If $(g_m : m \in \mathbb{N})$ is pointwise increasing and $\lim_{m\to\infty} g_m = g$ pointwise, then the sequence $\left(\int_\Omega g_m \mathrm{d}\mu : m \in \mathbb{N}\right)$ is increasing and $\lim_{m\to\infty} \int_\Omega g_m \mathrm{d}\mu = \int_\Omega g \mathrm{d}\mu$.*
3. *For some $\alpha, \beta \in [0, \infty]$, we have*

$$\int_\Omega \alpha g + \beta h \mathrm{d}\mu = \alpha \int_\Omega g \mathrm{d}\mu + \beta \int_\Omega h \mathrm{d}\mu.$$

(We use the convention "$0 \cdot \infty = 0$")

Proof. Lemma 4.6 in [25]. □

Measurable functions g taking values in $\mathbb{R}$ can be integrated by subtracting the integral of their negative part $\max\{0, -g\}$ from the integral of their positive part $\max\{0, g\}$, if one of them is integrable.

Integrals of non-negative measurable functions give a natural way to define measures.

Proposition and definition 6.2.12. *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and let $g : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}\mathbb{R})$ be measurable and non-negative. Then, the map $\nu : \mathcal{F} \to [0, \infty]$, defined by*

$$F \mapsto \int_\Omega g \cdot \mathbf{1}_F \mathrm{d}\mu =: \int_F g \mathrm{d}\mu$$

is a measure. ν is called measure with (μ-)density (function) g. If ν is a probability measure, g is called (μ-)probability density (function).

Proof. Exercise. □

Definition 6.2.13. *Let $(\Omega, \mathcal{F}, \mu) := (\mathbb{R}, \mathcal{B}\mathbb{R}, \lambda_1)$. Moreover, let $m \in \mathbb{R}$ and $\sigma > 0$, and let $g : \Omega \to \mathbb{R}$ be the measurable function*

$$g(\omega) := \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(\omega - m)^2}{2\sigma^2}\right).$$

Then, the measure ν with λ_1-density g is called Gaussian distribution on $\mathbb{R}$ with mean m and variance σ^2. We denote $\mathrm{n}(\cdot; m, \sigma^2) := g$ and $\mathrm{N}(m, \sigma^2) := \nu$. Moreover, we define the degenerate Gaussian distribution by $\mathrm{N}(m, 0) := \delta(\cdot - m)$.

A rather surprising result about measures and densities is the Radon–Nikodym Theorem. It is fundamental for the general definition of conditional expectations and also for the general form of Bayes' theorem. Before stating the Radon–Nikodym Theorem, we define two more important notions regarding measures.

Definition 6.2.14. *Let $(\Omega, \mathcal{F})$ be a measurable space and μ, ν be two measure on that space.*

1. *We define ν to be absolutely continuous with respect to μ, if for all $F \in \mathcal{F}$, with $\mu(F) = 0$, we also have $\nu(F) = 0$. In this case, we write $\nu \ll \mu$.*
2. *Let $A(\omega)$ be a statement for all $\omega \in \Omega$. We say that A holds μ-almost everywhere (μ-a.e.), if there is a set $N \in \mathcal{F}$ such that $\mu(N) = 0$ and $A(\omega)$ is true for $\omega \in X \backslash N$. If μ is a probability measure, we sometimes say μ-almost surely (μ-a.s.) instead of μ-almost everywhere.*

Theorem 6.2.15 (Radon-Nikodym)**.** *Let $(\Omega, \mathcal{F})$ be a measurable space and let μ, ν be σ-finite measures on $(\Omega, \mathcal{F})$. Then, the following two statements are equivalent:*

(i) $\nu \ll \mu$

(ii) There is a measurable function $g : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}\mathbb{R})$, with

$$\nu(F) = \int_F g \mathrm{d}\mu \qquad (F \in \mathcal{F}).$$

Moreover, the function g is μ-a.e. unique, called Radon–Nikodym derivative, and denoted by $\frac{\mathrm{d}\nu}{\mathrm{d}\mu} := g$.

Proof. (ii) ⇒ (i): exercise. (i) ⇒ (ii): more complicated, see, e.g., Corollary 7.34 in [25]. □

Exercise 6.2.16. Give an example for measures ν, μ on $(\mathbb{R}, \mathcal{B}\mathbb{R})$, with $\nu \ll \mu$ and μ not σ-finite, such that no Radon-Nikodym derivative exists.

6.3 Conditional probability

For the remainder of the lecture, we always consider $(\Omega, \mathcal{F}, \mathbb{P})$ as underlying probability space for any random variable. We typically omit its precise construction, but assume that Ω is a Polish space (separable and completely metrisable) and $\mathcal{F} := \mathcal{B}\Omega$. We denote integrals with respect to $\mathbb{P}$ sometimes by

$$\mathbb{E}[\varphi] := \int_\Omega \varphi \mathrm{d}\mathbb{P},$$

for some $\varphi : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}\mathbb{R})$, for which this integral is well-defined.

Example 6.3.1. Let $U : (\Omega, \mathcal{F}) \to (\{1, \dots, 6\}, 2^{\{1,\dots,6\}})$ be a random variable modelling the roll of a die, hence

$$\mathbb{P}(U = u) = \begin{cases} 1/6, & \text{if } u \in \{1, ..., 6\}, \\ 0, & \text{otherwise.} \end{cases}$$

This probability measure models our knowledge concerning the outcome of the experiment. Now we consider an extended model. After the die is rolled and before its realisation is revealed, we are told whether the realisation is even or odd. Given this information, we can adjust our knowledge concerning the random variable U:

$$\mathbb{P}(U = u | U \text{ is even}) = \frac{\mathbb{P}(U = u \text{ and } U \text{ is even})}{\mathbb{P}(U \text{ is even})},$$

respectively

$$\mathbb{P}(U = u | U \text{ is odd}) = \frac{\mathbb{P}(U = u \text{ and } U \text{ is odd})}{\mathbb{P}(U \text{ is odd})}.$$

In the example above, we used the elementary definition of conditional probabilities:

$$\mathbb{P}(F|F') = \frac{\mathbb{P}(F \cap F')}{\mathbb{P}(F')} \qquad (F, F' \in \mathcal{F}, \mathbb{P}(F') > 0).$$

This definition can only be used, if the event with respect to which the conditional probability is defined has a positive probability (here: $\{U \text{ is even}\}, \{U \text{ is odd}\}$).

This however is typically not the case in a Bayesian inverse problem since the probability measure of the noise is continuous. Hence, we need a more general definition of conditional probabilities. We start with conditional expectations.

Theorem 6.3.2. *Let $U : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}\mathbb{R})$ and $Y : (\Omega, \mathcal{F}) \to (\mathcal{Y}, \mathcal{B}\mathcal{Y})$ be random variables and let U be integrable. Then, there exists a measurable function $h : (\mathcal{Y}, \mathcal{B}\mathcal{Y}) \to (\mathbb{R}, \mathcal{B}\mathbb{R})$, such that*

$$\int_F h(y)\mathbb{P}(Y \in \mathrm{d}y) = \int_{\{Y \in F\}} U \mathrm{d}\mathbb{P} \qquad (F \in \mathcal{B}\mathcal{Y}). \tag{6.1}$$

Moreover, h is $\mathbb{P}(Y \in \cdot)$-a.s. unique.

Proof. We assume without loss of generality that $U \geqslant 0$. (If U is real-valued, study $\max\{U, 0\}$ and $\max\{-U, 0\}$ separately.) Note that the map

$$F \mapsto \int_{\{Y \in F\}} U \mathrm{d}\mathbb{P} =: \mu(F)$$

defines a (σ-)finite measure. We now show that $\mu \ll \mathbb{P}(Y \in \cdot)$: let $F_0 \in \mathcal{B}\mathcal{Y}$ be chosen such that $\mathbb{P}(Y \in F_0) = 0$. Then,

$$\int_{\{Y \in F_0\}} U \mathrm{d}\mathbb{P} = \int_\Omega \mathbf{1}_{\{Y \in F_0\}} U \mathrm{d}\mathbb{P} = 0.$$

By the Radon–Nikodym Theorem, there exists a $\mathbb{P}(Y \in \cdot)$-a.s. unique function $h := \frac{\mathrm{d}\mu}{\mathrm{d}\mathbb{P}(Y \in \cdot)}$, satisfying (6.1). □

Definition 6.3.3. *$h(y)$ in Theorem 6.3.2 is called conditional expectation of U given $Y = y$. We write $h(y) =: \mathbb{E}[U|Y = y]$, for $\mathbb{P}(Y \in \cdot)$-almost every $y \in \mathcal{Y}$.*

Now we can define the conditional probability of some event F by considering the indicator random variable $U = \mathbf{1}_F$. Since $\mathcal{X}, \mathcal{Y}$ are Polish spaces, one can even find a $\mathbb{P}(Y \in \cdot)$-a.s. unique Markov kernel $(y, F) \mapsto \mathbb{E}[\mathbf{1}_F|Y = y]$.

Definition 6.3.4. *Let $(\Omega, \mathcal{F}), (\Omega', \mathcal{F}')$ be measurable spaces. A map $M : \Omega \times \mathcal{F}' \to [0, 1]$ is called Markov kernel from $(\Omega, \mathcal{F})$ to $(\Omega', \mathcal{F}')$, if*

(i) $M(\omega, \cdot)$ is a probability measure for all $\omega \in \Omega$,

(ii) $M(\cdot, F') : (\Omega, \mathcal{F}) \to ([0, 1], \mathcal{B}[0, 1])$ is measurable for all $F' \in \mathcal{F}'$.

Theorem 6.3.5. *Let $U : (\Omega, \mathcal{F}) \to (\mathcal{X}, \mathcal{B}\mathcal{X})$ and $Y : (\Omega, \mathcal{F}) \to (\mathcal{Y}, \mathcal{B}\mathcal{Y})$ be random variables. Then, there exist a Markov kernel M from $(\mathcal{Y}, \mathcal{B}\mathcal{Y})$ to $(\mathcal{X}, \mathcal{B}\mathcal{X})$, with*

$$\int_F M(y, F')\mathbb{P}(Y \in \mathrm{d}y) = \mathbb{P}(\{Y \in F\} \cap \{U \in F'\}) \qquad (F \in \mathcal{B}\mathcal{Y}, F' \in \mathcal{B}\mathcal{X}).$$

Moreover, M is $\mathbb{P}(Y \in \cdot)$-a.s. unique.

Proof. Non-trivial, but possible if Ω is Polish; see [26]. □

Definition 6.3.6. *M in Theorem 6.3.5 is called (regular) conditional probability distribution of U given $Y = y$. We write $M(y, F) := \mathbb{P}(U \in F|Y = y)$, for $F \in \mathcal{B}\mathcal{X}, y \in \mathcal{Y}$.*

Example 6.3.7 (Example 6.3.1 rev.)**.** In Example 6.3.1, we compute the conditional probability distribution of a die $U : (\Omega, \mathcal{F}) \to (\{1, \ldots, 6\}, 2^{\{1,\ldots,6\}})$, given the information whether the outcome will be even or odd. Define a random variable $Y : (\Omega, \mathcal{F}) \to (\{0, 1\}, 2^{\{0,1\}})$

$$\omega \mapsto \begin{cases} 0, & \text{if } U(\omega) \text{ is even} \\ 1, & \text{otherwise.} \end{cases}$$

We can write

$$\mathbb{P}(U = u|U \text{ is even}) =: \mathbb{P}(U = u|Y = 0), \quad \mathbb{P}(U = u|U \text{ is odd}) =: \mathbb{P}(U = u|Y = 1).$$

Indeed, one can show that these functions are conditional expectation/probability measures in the sense of definition 6.3.3. Let $F \in 2^{\{0,1\}}$. We need to show that

$$\int_F \mathbb{P}(U = u|Y = y)\mathbb{P}(Y \in \mathrm{d}y) = \mathbb{P}(\{U = u\} \cap \{Y \in F\}).$$

Let $F := \{0\}$. Then, we have

$$\begin{aligned}\int_{\{Y=0\}} \mathbf{1}_{\{U=u\}} \mathrm{d}\mathbb{P} &= \frac{1}{6}(\mathbf{1}_{\{2\}}(u) + \mathbf{1}_{\{4\}}(u) + \mathbf{1}_{\{6\}}(u)) \\ &= \underbrace{\frac{1}{2}}_{=\mathbb{P}(Y=0)} \cdot \underbrace{\frac{1}{3}(\mathbf{1}_{\{2\}}(u) + \mathbf{1}_{\{4\}}(u) + \mathbf{1}_{\{6\}}(u))}_{=\mathbb{P}(U=u|Y=0)} \\ &= \int_{\{0\}} \mathbb{P}(U = u|Y = y)\mathbb{P}(Y \in \mathrm{d}y)\end{aligned}$$

Analogously, one can show condition (6.1) for $F = \emptyset, \{1\}, \{0, 1\}$.

In Theorem 6.3.5, we discuss that conditional probabilities are Markov kernels. Also the converse is true: given a Markov kernel, we can construct random variables such that the Markov kernel represents a conditional probability measure.

Proposition 6.3.8. *Let $M : \Omega' \times \mathcal{F}'' \to [0, 1]$ be a Markov kernel from $(\Omega', \mathcal{F}')$ to $(\Omega'', \mathcal{F}'')$. Then, there is an underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and random variables $X' : \Omega \to \Omega'$ and $X'' : \Omega \to \Omega''$ such that:*

$$M(\omega', F'') = \mathbb{P}(X'' \in F''|X' = \omega') \quad (F'' \in \mathcal{F}'' \text{ and } \mathbb{P}(X' \in \cdot)\text{-almost all } \omega' \in \Omega').$$

Proof. Define $(\Omega, \mathcal{F}) := (\Omega' \times \Omega'', \mathcal{F}' \otimes \mathcal{F}'')$. Let μ' be some probability measure on $(\Omega', \mathcal{F}')$. Moreover, let $\mathbb{P}$ be the measure satisfying

$$\mathbb{P}(F' \times F'') = \int_{F'} M(\omega', F'')\mathrm{d}\mu'(\omega') \quad (F' \in \mathcal{F}', F'' \in \mathcal{F}'').$$

Let $X' : \Omega \to \Omega'$ (resp. $X'' : \Omega \to \Omega''$) be the canonical projection on the first (resp. second) coordinate. Then $X' \sim \mu'$ and $X'' \sim M(X', \cdot)$. Let $F' \in \mathcal{F}'$ and $F'' \in \mathcal{F}''$. Then it holds

$$\begin{aligned}\mathbb{P}(\{X' \in F'\} \cap \{X'' \in F''\}) = \int_{\{X' \in F', X'' \in F''\}} \mathrm{d}\mathbb{P} &= \iint_{\{X' \in F', X'' \in F''\}} M(\omega', \mathrm{d}\omega'')\mu'(\mathrm{d}\omega') \\ &\overset{(*)}{=} \int_{F'} \int_{F''} M(\omega', \mathrm{d}\omega'')\mu'(\mathrm{d}\omega') = \int_{F'} M(F'', \omega')\mathbb{P}(X' \in \mathrm{d}\omega'),\end{aligned}$$

where (*) is implied by Tonelli's Theorem. Hence, $M(F'', \omega') = \mathbb{P}(X'' \in F''|X' = \omega')$ is indeed a conditional probability distribution. □

As Markov kernels are consistent with conditional probabilities, we sometimes write $M(\cdot|*) := M(*, \cdot)$.

Applying the concept of conditional expectations in general situations is not straightforward. However, probability measures are often given in terms of probability density functions. Given joint and marginal probability density functions, one can define the conditional probability in terms of a probability density function.

Lemma 6.3.9. *Let U, Y be random variables with joint probability distribution $\mathbb{P}((U, Y) \in \cdot)$ that is absolutely continuous with respect to a σ-finite measure ν on $(\mathcal{X} \times \mathcal{Y}, \mathcal{B}\mathcal{X} \otimes \mathcal{B}\mathcal{Y})$. Assume that $\nu = \nu_U \otimes \nu_Y$ for σ-finite measure spaces $(\mathcal{X}, \mathcal{B}\mathcal{X}, \nu_U)$, $(\mathcal{Y}, \mathcal{B}\mathcal{Y}, \nu_Y)$. We write $g_{U,Y} := \frac{\mathrm{d}\mathbb{P}((U,Y)\in\cdot)}{\mathrm{d}\nu}$ for the joint probability density function. Then,*

$$\mathbb{P}(U \in \cdot) \ll \nu_U, \quad \mathbb{P}(Y \in \cdot) \ll \nu_Y,$$

with probability density functions

$$g_U := \int_{\mathcal{Y}} g_{U,Y} \mathrm{d}\nu_Y = \frac{\mathrm{d}\mathbb{P}(U \in \cdot)}{\mathrm{d}\nu_U} \quad (\nu_U\text{-a.e.}),$$
$$g_Y := \int_{\mathcal{X}} g_{U,Y} \mathrm{d}\nu_U = \frac{\mathrm{d}\mathbb{P}(Y \in \cdot)}{\mathrm{d}\nu_Y} \quad (\nu_Y\text{-a.e.}).$$

Proof. Let $A \in \mathcal{B}\mathcal{X}$. We have

$$\mathbb{P}(U \in A) = \mathbb{P}(U \in A, Y \in \mathcal{Y}) = \int_{A\times\mathcal{Y}} g_{U,Y} \mathrm{d}\nu = \int_A \underbrace{\int_{\mathcal{Y}} g_{U,Y} \mathrm{d}\nu_Y}_{=:g_U} \mathrm{d}\nu_U,$$

by the Theorem of Tonelli. Hence, $\mathbb{P}(U \in \cdot) \ll \nu_U$. The statement about Y can be proven analoguously. □

Theorem 6.3.10. *Under the assumptions of Lemma 6.3.9, we have $\mathbb{P}(U \in \cdot|Y = y) \ll \nu_U$ with ν_U-density:*

$$g_{U|Y=y}(u) := \begin{cases} \frac{g_{U,Y}(u,y)}{g_Y(y)}, & \text{if } g_Y(y) > 0, \\ 0, & \text{otherwise} \end{cases} \quad (u \in \mathcal{X}, \nu_U\text{-a.e.}; y \in \mathcal{Y}, \mathbb{P}(Y \in \cdot)\text{-a.e.}),$$

and equivalently $\mathbb{P}(Y \in \cdot|U = u) \ll \nu_Y$ with ν_Y-density:

$$g_{Y|U=u}(y) := \begin{cases} \frac{g_{U,Y}(u,y)}{g_U(u)}, & \text{if } g_U(u) > 0, \\ 0, & \text{otherwise} \end{cases} \quad (y \in \mathcal{Y}, \nu_Y\text{-a.e.}; u \in \mathcal{X}, \mathbb{P}(U \in \cdot)\text{-a.e.}).$$

Proof. Let $A \in \mathcal{B}\mathcal{X}, F \in \mathcal{B}\mathcal{Y}$. By Definition 6.3.6, $\mathbb{P}(U \in A|Y = y)$ fulfills (6.1):

$$\begin{aligned} \mathbb{P}(U \in A, Y \in F) &= \int_F \mathbb{P}(U \in A|Y = y)\mathbb{P}(Y \in \mathrm{d}y) \\ &= \int_F \mathbb{P}(U \in A|Y = y) g_Y(y) \mathrm{d}\nu_Y(y) \\ &= \int_{F\cap\{g_Y>0\}} \mathbb{P}(U \in A|Y = y) g_Y(y) \mathrm{d}\nu_Y(y), \end{aligned}$$

as $\mathbb{P}(g_Y(Y) = 0) = \mathbb{P}(Y \in \{g_Y = 0\}) = \int_{\mathcal{Y}} \mathbf{1}_{\{g_Y=0\}} \mathbb{P}(Y \in \mathrm{d}y) = \int_{\{g_Y=0\}} g_Y \mathrm{d}\nu_Y = 0$. Note that we can write

$$\mathbb{P}(U \in A, Y \in F) = \int_{F\cap\{g_Y>0\}} \int_A g_{U,Y}(u,y) \mathrm{d}\nu_U(u) \mathrm{d}\nu_Y(y).$$

This and the statement above imply

$$\mathbb{P}(U \in A|Y = y) g_Y(y) = \int_A g_{U,Y}(u,y) \mathrm{d}\nu_U(u) \qquad (\mathbb{P}(Y \in \cdot)\text{-a.s.}).$$

Hence, we have

$$\mathbb{P}(U \in A|Y = y) = \int_A \frac{g_{U,Y}(u,y)}{g_Y(y)} \mathrm{d}\nu_U(u) \qquad (\mathbb{P}(Y \in \cdot)\text{-a.s.}).$$

This proves our statement about $\mathbb{P}(U \in \cdot|Y = y)$ the reverse statement can be shown analoguously. □

Definition 6.3.11. *Let $g_U, g_Y, g_{U,Y}, g_{U|Y=y}, g_{Y|U=u}$ be the probability density functions in Theorem 6.3.10. We define*

- *g_U (resp. g_Y) to be the marginal probability density of U (resp. Y),*
- *$g_{U,Y}$ to be the joint probability density of U and Y,*
- *$g_{U|Y=y}$ to be the conditional density of U given $Y = y$, and*
- *$g_{Y|U=u}$ to be the conditional density of Y given $U = u$.*

6.4 Bayesian statistics

We are now ready to, first, fit our inverse problem into a statistical framework and, second, determine the posterior measure

6.4.1 Statistical models

Definition 6.4.1. *Let $\mathcal{X}, \mathcal{Y}$ be separable Banach spaces. We refer to $\mathcal{X}$ as parameter space and to $\mathcal{Y}$ as data space. Let $\mathcal{P} := \{M(\cdot|u) : u \in \mathcal{X}\}$, where M is a Markov kernel from $(\mathcal{X}, \mathcal{B}\mathcal{X})$ to $(\mathcal{Y}, \mathcal{B}\mathcal{Y})$. The tuple $(\mathcal{Y}, \mathcal{P})$ is called statistical model. The statistical model is called parametric, if $\mathcal{X}$ is a subset of a Euclidean vector space, and non-parametric, otherwise.*

After defining statistical models, we should comment on their purpose.

Remark 6.4.2. Let $u^* \in \mathcal{X}$ be some parameter, let $Y \sim M(\cdot|u^*)$, and let y be a realisation of Y. Statistical methods aim to find $u^* \in \mathcal{X}$ based on the realisation y. The probability measure $M(\cdot|u^*)$ is called data-generating distribution.

Now, we give an example for a parametric statistical model.

Example 6.4.3. We are given five independent realisations $y = (0.2, -0.32, 0.8, 1.2, -0.4)$, of a one dimensional Gaussian random variable with variance $\sigma^2 = 1$. We do not know the mean of the random variable. Given y, we want to identify the mean. The statistical model associated with this task is given by:

$$(\mathcal{Y}, \mathcal{P}) := \left(\mathbb{R}^5, \left\{\mathrm{N}(u, 1)^{\otimes 5} : u \in \mathbb{R}\right\}\right).$$

We can sometimes represent a statistical model in terms of a conditional density, the so-called likelihood.

Definition 6.4.4. *Let $(\mathcal{Y}, \mathcal{P})$ be a statistical model and let $L : (\mathcal{X} \times \mathcal{Y}, \mathcal{B}\mathcal{X} \otimes \mathcal{B}\mathcal{Y}) \to (\mathbb{R}, \mathcal{B}\mathbb{R})$ such that*

$$\mathcal{P} := \left\{\mathcal{B}\mathcal{Y} \ni F \mapsto \int_F L(y|u)\mathrm{d}\mu(y) : u \in \mathcal{X}\right\},$$

for some measure μ on $(\mathcal{Y}, \mathcal{B}\mathcal{Y})$. We refer to L as (data) likelihood.

Note that the likelihood is a conditional density $L = g_{Y|U=u}$, for some random variable U. It informs us about the likelihood of observing a data set given that we assume it was sampled from $M(\cdot|u)$.

Example 6.4.5. Let $\mathcal{A} : (\mathcal{X}, \mathcal{B}\mathcal{X}) \to (\mathcal{Y}, \mathcal{B}\mathcal{Y})$ be a measurable operator. Moreover, let μ_{noise} be a probability measure on $(\mathcal{Y}, \mathcal{B}\mathcal{Y})$. We consider the inverse problem of identifying $u \in X$, where

$$\mathcal{A}(u) + N = f_n$$

with $N \sim \mu_{\text{noise}}$. We can now represent this inverse problem by a statistical model:

$$(\mathcal{Y}, \mathcal{P}) := (\mathcal{Y}, \{\mu_{\text{noise}}(\cdot - \mathcal{A}(u)) : u \in \mathcal{X}\}).$$

The data set f_n is a realisation of the data-generating distribution $\mu_{\text{noise}}(\cdot - \mathcal{A}(u^*))$, where u^* is the true parameter.

Let $n \in \mathbb{N}$, $\mathcal{Y} := \mathbb{R}^n$, $\Gamma \in \mathbb{R}^{n\times n}$ be positive definite, and $\mu_{\text{noise}} := \mathrm{N}(0, \Gamma)$. Then, we can represent the statistical model by a likelihood:

$$L(y|u) := (2\pi)^{-k/2}\det(\Gamma)^{-1/2} \exp\left(-\frac{1}{2}\|\Gamma^{-1/2}(y - \mathcal{A}(u))\|^2\right),$$

where $u \in \mathcal{X}$ and $y \in \mathcal{Y}$.

6.4.2 Bayes' formula

In Bayesian statistics, we model the unknown parameter u as a random variable $U \sim \mu_0$ that is distributed according to a prior measure. μ_0 reflects our knowledge concerning the parameter u before seeing the data. Moreover, we are given a statistical model $(\mathcal{Y}, \mathcal{P})$ and the according Likelihood L, which is a conditional density $f_{Y|U=u}$. We aim to *invert* $\mathbb{P}(Y \in \cdot|U = \cdot)$ to $\mathbb{P}(U \in \cdot|Y = \cdot)$. The conditional measure $\mathbb{P}(U \in \cdot|Y = \cdot)$ is the updated prior $\mathbb{P}(U \in \cdot) := \mu_0$. This updating/inversion process uses on Bayes' formula.

Theorem 6.4.6 (Bayes)**.** *Let U, Y be random variables as in Theorem 6.3.10. Then,*

$$g_{U|Y=y}(u) = \frac{g_{Y|U=u}(y)g_U(u)}{g_Y(y)}, \tag{6.2}$$

for $u \in \mathcal{X}$, ν_U-a.e. and $y \in \mathcal{Y}$, $\mathbb{P}(Y \in \cdot)$-a.e. with $g_Y(y) > 0$.

Proof. We need to show that $g_{Y|U=u}g_U = g_{U,Y}$, $\nu_U \otimes \mathbb{P}(Y \in \cdot)$-a.e.. Let $u \in \mathcal{X}$ with $g_U(u) > 0$. By definition,

$$g_{Y|U=u}(y)g_U(u) = \frac{g_{U,Y}(u, y)g_U(u)}{g_U(u)} = g_{U,Y}(u, y) \qquad ((u, y) \in \{g_U > 0\} \times \mathcal{Y}, \nu_U \otimes \mathbb{P}(Y \in \cdot)\text{-a.e.}).$$

Conversely, let $u \in \mathcal{X}$, with $g_U(u) = 0$. This implies that

$$0 = \int_{\mathcal{Y}} g(u, y)\mathrm{d}\nu_Y(y).$$

Then, $g_{U,Y}(u, \cdot) = 0$, ν_Y-a.e. and, thus, also $\mathbb{P}(Y \in \cdot)$-a.s.. Hence, $g_{U,Y} = 0 = g_{Y|U=u}g_U$. □

Definition 6.4.7.

- *$Z(y) := g_Y(y)$ is called (model) evidence or marginal likelihood*[1]*,*
- *$L(y|u) := g_{Y|U=u}(y)$ is called (data) likelihood,*

[1] $Z(y)$ is derived from German: *Zustandssumme ('sum of states')*

- $\mu_0 := \mathbb{P}(U \in \cdot)$ *is called* prior (measure),
- $\mu_{\text{post}} := \mathbb{P}(U \in \cdot | Y = y)$ *is called* posterior (measure), *and*

In Theorem 6.4.6, we require that μ_0 has a probability density function g_U with respect to a measure ν_U. In practice, ν_U is often a Lebesgue measure or the counting measure. In some cases, neither of those two is well-defined or a sensible choice, e.g., when $\dim \mathcal{X} = \infty$. However, we can always assume that $\nu_U := \mu_0$. In this case, we obtain the formulation of Stuart [35]:

$$\frac{\mathrm{d}\mu_{\text{post}}}{\mathrm{d}\mu_0}(u) = \frac{L(y|u)}{Z(y)} \quad (u \in \mathcal{X}, \mu_0\text{-a.s.}).$$

Remark 6.4.8. When defining $Z(y) := \int L(y|u)\mathrm{d}\mu_0$, it is not necessary for $L(y|u)$ to be correctly normalised. Indeed, we can set $L(y|u) := c \cdot g_{Y|U=u}(y)$, for some constant $c > 0$. The factor c cancels with the same factor in $Z(y)$. However, then we have $Z(y) \neq f_Y(y)$ and call $Z(y)$ normalising constant.

Chapter 7

Bayesian inverse problems and well-posedness

In this chapter, we will define Bayesian inverse problems and study their well-posedness. Well-posedness requires existence and uniqueness of the posterior measure, as well as its stability with respect to marginal perturbations in the data.

7.1 Bayesian inverse problems

A posterior measure is a conditional probability distribution and as such only for almost every data set uniquely defined. In the following, we will always pick one representing Markov kernel out of the set of kernels satisfying the equation in Theorem 6.3.5. We do so, by fixing the definition of the likelihood to a specific measurable function $\mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ and defining the posterior to satisfy Bayes' formula with this likelihood.

We first introduce some further notation.

Definition 7.1.1. *Let $(\Omega', \mathcal{F}')$ be some measurable space. We define the space of probability measures on $(\Omega', \mathcal{F}')$ by $\mathrm{Prob}(\Omega', \mathcal{F}') := \{\mu : \mu \text{ is a probability measure on } (\Omega', \mathcal{F}')\}$. Moreover, for some σ-finite measure ν on $(\Omega', \mathcal{F}')$, we define $\mathrm{Prob}(\Omega', \mathcal{F}', \nu) := \{\mu \in \mathrm{Prob}(\Omega, \mathcal{F}') : \mu \ll \nu\}$.*

Definition 7.1.2. *Let $\mu_0 \in \mathrm{Prob}(\mathcal{X}, \mathcal{B}\mathcal{X})$ and $L : (\mathcal{X} \times \mathcal{Y}, \mathcal{B}\mathcal{X} \otimes \mathcal{B}\mathcal{Y}) \to (\mathbb{R}, \mathcal{B}\mathbb{R})$ be a measureable function. We define the Bayesian inverse problem (BIP) with prior μ_0 and likelihood L, to be the problem of finding $\mu_{\mathrm{post}} \in \mathrm{Prob}(\mathcal{X}, \mathcal{B}\mathcal{X})$ with*

$$\frac{\mathrm{d}\mu_{\mathrm{post}}}{\mathrm{d}\mu_0}(u) = \frac{L(f_n|u)}{\int_{\mathcal{X}} L(f_n|u)\mathrm{d}\mu_0(u)} \qquad (u \in \mathcal{X}; \mu_0\text{-a.s.})$$

for any data set $f_n \in \mathcal{Y}$.

We discussed previously how to construct a likelihood in the 'classical' inverse problem setting

$$\text{find } u \in \mathcal{X} : \mathcal{A}(u) + n = f_n.$$

We now allow for much more general likelihood functions; this includes non-additive noise, Poissonian models,...

Definition 7.1.3. *Consider a (BIP) with prior μ_0 and likelihood L. Let $P \subseteq \mathrm{Prob}(\mathcal{X}, \mathcal{B}\mathcal{X})$ be a space of probability measures and $d : P^2 \to [0, \infty)$ be a metric on P. A Bayesian inverse problem is $\underline{(P,d)\text{-well-posed}}$, if*

(i) for all $f_n \in \mathcal{Y}$, the probability measure $\mu_{\mathrm{post}} \in P$ exists, (existence)

(ii) for all $f_n \in \mathcal{Y}$, the probability measure $\mu_{\mathrm{post}} \in P$ is unique, and (uniqueness)

(iii) the map $\mathcal{Y} \ni f_n \mapsto \mu_{\mathrm{post}} \in P$ is continuous. (stability)

Existence and uniqueness of the posterior in $P \in \{\mathrm{Prob}(\mathcal{X}, \mathcal{B}\mathcal{X}), \mathrm{Prob}(\mathcal{X}, \mathcal{B}\mathcal{X}, \mu_0)\}$ is automatic, if $\int_{\mathcal{X}} L(f_n|u)\mathrm{d}\mu_0(u) \in (0, \infty)$. This is, for instance, the case in the following lemma.

Lemma 7.1.4. *Consider a (BIP) with prior μ_0 and likelihood L. Let $L > 0$ (μ_0-a.s.) and $L(f_n|\cdot) \in L^1(\mathcal{X}, \mathcal{B}\mathcal{X}, \mu_0)$ for any $f_n \in \mathcal{Y}$. Then, the posterior $\mu_{\mathrm{post}} \in \mathrm{Prob}(\mathcal{X}, \mathcal{B}\mathcal{X}, \mu_0)$ exists and is unique.*

Proof. We need to show that $\int_{\mathcal{X}} L(f_n|u)\mathrm{d}\mu_0(u) \in (0, \infty)$. Upper bound: trivial, since $L(f_n|\cdot) \in L^1(\mathcal{X}, \mathcal{B}\mathcal{X}, \mu_0)$. Lower bound: exercise. □

Before we can actually speak about stability, we need to discuss metrics on spaces of probability measures.

7.2 Metrics on spaces of probability measures

We consider metrics on subspaces of $\mathrm{Prob}(\mathcal{X}, \mathcal{B}\mathcal{X})$ to be able to show stability of the posterior measure with respect to perturbations in the data. We consider two different concept: total variation and weak convergence.

Definition 7.2.1. *(i) Let $(\Omega', \mathcal{F}')$ be a measurable space. We define the $\underline{\text{total variation (TV) distance}}$ on $\mathrm{Prob}(\Omega', \mathcal{F}')$ by*

$$d_{\mathrm{TV}} : \mathrm{Prob}(\Omega', \mathcal{F}')^2 \to [0, \infty), (\mu, \nu) \mapsto \sup_{F' \in \mathcal{F}'} |\mu(F') - \nu(F')|$$

(ii) Let Ω' be a topological space and $(\Omega', \mathcal{F}') := (\Omega', \mathcal{B}\Omega')$. Let $(\mu_n)_{n\in\mathbb{N}} \in \mathrm{Prob}(\Omega', \mathcal{F}')^{\mathbb{N}}$ and $\mu \in \mathrm{Prob}(\Omega', \mathcal{F}')$. We say $\mu_n \to \mu$ $\underline{\text{weakly}}$, as $n \to \infty$, if

$$\lim_{n\to\infty} \int_{\Omega'} g\mathrm{d}\mu_n = \int_{\Omega'} g\mathrm{d}\mu,$$

for any $g : (\Omega', \mathcal{B}\Omega') \to (\mathbb{R}, \mathcal{B}\mathbb{R})$ that is continuous and bounded.

Remark 7.2.2. Weak convergence of measures on $\mathrm{Prob}(\mathcal{X}, \mathcal{B}\mathcal{X})$ can be represented by the (Lévy)-Prokhorov metric d_{LP}. See [29] for details. Hence, when referring to the topology induced by weak convergence, we will usually speak about the metric space $(\mathrm{Prob}(\mathcal{X}, \mathcal{B}\mathcal{X}), d_{\mathrm{LP}})$, but not actually employ the (Lévy)-Prokhorov metric.

We end this section with two more results about the total variation distance and weak convergence. First, we show that if a sequence of measures converges in the total variation distance, it converges weakly as well.

Lemma 7.2.3. *Let Ω' be a topological space and $(\Omega', \mathcal{F}') := (\Omega', \mathcal{B}\Omega')$. Let $(\mu_n)_{n\in\mathbb{N}} \in \mathrm{Prob}(\Omega', \mathcal{F}')^{\mathbb{N}}$ and $\mu \in \mathrm{Prob}(\Omega', \mathcal{F}')$. Then*

$$\lim_{n\to\infty} d_{\mathrm{TV}}(\mu_n, \mu) = 0 \Longrightarrow \mu_n \to \mu, \text{ weakly as } n \to \infty.$$

The converse statement ("$\Leftarrow$") is in general not true.

Proof. Exercise. □

Second, we give a representation of the total variation distance of two measures having a density with respect to a third measure.

Lemma 7.2.4. *Let $\mu, \nu \in \mathrm{Prob}(\Omega, \mathcal{F})$ and ρ be a σ-finite measure with $\mu, \nu \ll \rho$. Then, $d_{\mathrm{TV}}(\mu, \nu) = \frac{1}{2}\int_\Omega \left|\frac{\mathrm{d}\mu}{\mathrm{d}\rho} - \frac{\mathrm{d}\nu}{\mathrm{d}\rho}\right| \mathrm{d}\rho$.*

Proof. Exercise. □

Note that this result is independent of the measure ρ. As a trivial dominating measure, one can always choose $\rho := \mu + \nu$.

7.3 Stability

We now give a set of assumptions under which we can prove (P, d)-well-posedness, as defined in Definition 7.1.3, where (P, d) refers to the space of probability measure on $\mathcal{X}$ with μ_0-density and either total variation distance or weak convergence.

Assumption 7.3.1. Given a (BIP) with prior μ_0 and likelihood L. Let the following assumptions hold for $u \in \mathcal{X}$ μ_0-a.s. and $f_n \in \mathcal{Y}$.

(A1) $L(\cdot|u)$ is a strictly positive probability density function,

(A2) $L(f_n|\cdot) \in L^1(\mathcal{X}, \mathcal{B}\mathcal{X}, \mu_0)$,

(A3) some $h \in L^1(\mathcal{X}, \mathcal{B}\mathcal{X}, \mu_0)$ exists, such that $L(f_n'|\cdot) \leqslant h$ for all $f_n' \in \mathcal{Y}$, and

(A4) $L(\cdot|u)$ is continuous.

We now briefly comment on the assumptions. (A1) and (A2) were already required in Lemma 7.1.4. In (A3) we now not only ask for boundedness of the integral of the likelihood, but for its uniform boundedness by an integrable function g. This is for instance the case, if the likelihood is bounded by a constant (as μ_0 is a probability measure). In (A4) we ask for continuity in the data. (Continuity in the parameter is not required!) In inverse problems, this is true for a large number of noise distributions.

Before proving well-posedness under Assumptions (A1)-(A4) we cite a fundamental measure-theoretic result which is needed for the proof.

Theorem 7.3.2 (Dominated Convergence Theorem (DCT; Lebesgue))**.** *Let $(\Omega', \mathcal{F}', \mu')$ be a measure space. Let $g, (g_m)_{m\in\mathbb{N}}, h$ be measurable functions $(\Omega', \mathcal{F}') \to (\mathbb{R}, \mathcal{B}\mathbb{R})$ and $h \in L^1(\Omega', \mathcal{F}', \mu')$. Moreover, let $|g_m| \leqslant h$ (μ'-a.e.) and $g_m \to g$, μ'-a.e. as $m \to \infty$. Then, $g, g_m \in L^1(\Omega', \mathcal{F}', \mu')$ and*

$$\lim_{m\to\infty} \int_{\Omega'} g_m \mathrm{d}\mu' = \int_{\Omega'} g \mathrm{d}\mu'.$$

Proof. Can be proved using monotonic convergence theorem (Proposition 6.2.11.2). See, e.g., Theorem 1.6.9 [6] for a proof using Fatou's Lemma. □

Remark 7.3.3. The DCT describes a case in which we are allowed to "exchange integral and limit". The statement reads

$$\lim_{m\to\infty}\int_{\Omega'} g_m \mathrm{d}\mu' = \int_{\Omega'}\lim_{n\to\infty} g_m \mathrm{d}\mu'.$$

Equivalently, we could say it describes cases in which the integral as a functional of the integrand is continuous.

Theorem 7.3.4 (Well-posedness)**.** *Given a (BIP) with prior μ_0 and likelihood L that satisfies Assumptions (A1)–(A4). Moreover, let $P = \mathrm{Prob}(\mathcal{X}, \mathcal{B}\mathcal{X}, \mu_0)$ and $d \in \{d_{\mathrm{TV}}, d_{\mathrm{LP}}\}$. Then, the (BIP) is (P, d)-well-posed.*

Proof. 1. Note that (A1), (A2) already imply existence and uniqueness by Lemma 7.1.4. In the remainder of the proof, we focus on showing continuity in the total variation distance. Continuity in weak convergence is then implied by Lemma 7.2.3. Indeed, we show that for all $f_n \in \mathcal{Y}$ and all $(f_n^{(m)})_{m\in\mathbb{N}} \in \mathcal{Y}^{\mathbb{N}}$, with $\lim_{m\to\infty} f_n^{(m)} = f_n$, we have

$$\int_{\mathcal{X}} \left| \frac{L(f_n|u)}{Z(f_n)} - \frac{L(f_n^{(m)}|u)}{Z(f_n^{(m)})} \right| \mathrm{d}\mu_0(u) \to 0 \qquad (m \to \infty).$$

where $Z(f_n) := \int_{\mathcal{X}} L(f_n|u)\mathrm{d}\mu_0(u)$. By Lemma 7.2.4, this implies continuity of the posterior measure in the total variation distance.

2. We first show that $\mathcal{Y} \ni f_n \mapsto L(f_n|\cdot) \in L^1(\mathcal{X}, \mathcal{B}\mathcal{X}, \mu_0)$ is continuous. Let $f_n \in \mathcal{Y}$ and $(f_n^{(m)})_{m\in\mathbb{N}} \in \mathcal{Y}^{\mathbb{N}}$, with $\lim_{m\to\infty} f_n^{(m)} = f_n$. Note that

$$\lim_{m\to\infty}\int_{\mathcal{X}} \left| L(f_n^{(m)}|u) - L(f_n|u) \right| \mathrm{d}\mu_0(u) = \int_{\mathcal{X}} \lim_{m\to\infty} \left| L(f_n^{(m)}|u) - L(f_n|u) \right| \mathrm{d}\mu_0(u),$$

due to the DCT since the integrand is bounded below by 0 and above by $2h \in L^1(\mathcal{X}, \mathcal{B}\mathcal{X}, \mu_0)$. Due to the continuity of $L(\cdot|u)$ (required in (A4)), we have then

$$\lim_{m\to\infty}\int_{\mathcal{X}} \left| L(f_n^{(m)}|u) - L(f_n|u) \right| \mathrm{d}\mu_0(u) = 0.$$

With the same argument, we can show that $f_n \mapsto Z_n(f_n)$ is continuous.

3. The rest of the proof is similar to showing continuity of the quotient of two continuous functions. Let $f_n \in \mathcal{Y}$ and $(f_n^{(m)})_{m\in\mathbb{N}} \in \mathcal{Y}^{\mathbb{N}}$, with $\lim_{m\to\infty} f_n^{(m)} = f_n$. Then

$$\begin{aligned}
&\int_{\mathcal{X}} \left| \frac{L(f_n|u)}{Z(f_n)} - \frac{L(f_n^{(m)}|u)}{Z(f_n^{(m)})} \right| \mathrm{d}\mu_0(u) \\
&\leqslant Z(f_n)^{-1} \underbrace{\int_{\mathcal{X}} \left| L(f_n^{(m)}|u) - L(f_n|u) \right| \mathrm{d}\mu_0(u)}_{\to 0 \ (m\to\infty)} + \int_{\mathcal{X}} L(f_n^{(m)}|u)\mathrm{d}\mu_0(u) \underbrace{|Z(f_n)^{-1} - Z(f_n^{(m)})^{-1}|}_{\to 0 \ (m\to\infty)}
\end{aligned}$$

and the terms that do not converge to 0 are bounded. □

To illustrate the generality of this result, we now study again the inverse problem from Example 6.4.5.

Corollary 7.3.5. Let $k \in \mathbb{N}$ and $(\mathcal{Y}, \mathcal{B}\mathcal{Y}) := (\mathbb{R}^k, \mathcal{B}\mathbb{R}^k)$ and $\Gamma \in \mathbb{R}^{k\times k}$ be symmetric, positive definite. Moreover, let $\mathcal{A} : (\mathcal{X}, \mathcal{B}\mathcal{X}) \to (\mathcal{Y}, \mathcal{B}\mathcal{Y})$ be some function. Consider the (BIP) with some prior $\mu_0 \in \mathrm{Prob}(\mathcal{X}, \mathcal{B}\mathcal{X})$ and likelihood

$$L(f_n|u) := (2\pi)^{-k/2}\det(\Gamma)^{-1/2} \exp\left(-\frac{1}{2}\|\Gamma^{-1/2}(f_n - \mathcal{A}(u))\|^2\right) \qquad (u \in \mathcal{X}, f_n \in \mathcal{Y})$$

Then, the (BIP) is (P, d)-well-posed, with $P = \mathrm{Prob}(\mathcal{X}, \mathcal{B}\mathcal{X}, \mu_0)$ and $d \in \{d_{\mathrm{TV}}, d_{\mathrm{LP}}\}$.

Proof. Follows trivially from Theorem 7.3.4. □

Chapter 8

Function space priors and Monte Carlo

In this last chapter, we would like to discuss two rather practical topics:

- In inverse problems, we often consider infinite-dimensional parameter spaces. While we have discussed the well-posedness of Bayesian inverse problems in infinite dimensional setting, it is not clear yet how, e.g., a prior probability measure on such a space can be defined. We will discuss Gaussian prior measures on function spaces, so-called Gaussian random fields. For a more thorough introduction, we refer to the book by Bogachev [11].
- In practical situations, we need to approximate the posterior (or integrals with respect to it) numerically. We will discuss Monte Carlo techniques that are suitable for Bayesian inverse problems. Again, for a more thorough discussion of certain aspects, we refer to Agapiou et al. [3], Cotter et al. [16], and Robert and Casella [30].

8.1 Gaussian measures

We have defined Gaussian measures on $(\mathbb{R}, \mathcal{B}\mathbb{R})$ in Definition 6.2.13. We now extend this definition to measurable spaces like $(\mathcal{X}, \mathcal{B}\mathcal{X})$, where $\mathcal{X}$ is a separable Banach space. In this section, we assume that all Banach and Hilbert spaces are with respect to $\mathbb{R}$.

Definition 8.1.1. *Let μ be a probability measure on $\mathrm{Prob}(\mathcal{X}, \mathcal{B}\mathcal{X})$ and let $U \sim \mu$. We call μ Gaussian, if for all $\ell \in X^*$, there exist $m \in \mathbb{R}, \sigma \geqslant 0$, such that*

$$\mathbb{P}(\langle \ell, U \rangle \in \cdot) = \mathrm{N}(m, \sigma^2).$$

Moreover, we define the mean of μ by $a_\mu \in \mathcal{X}^{**}$, given by

$$a_\mu(\ell) = \int_{\mathcal{X}} \langle \ell, u \rangle \mathrm{d}\mu(u) \qquad (\ell \in \mathcal{X}^*)$$

and the covariance operator of μ by $R_\mu : \mathcal{X}^* \to \mathcal{X}^{**}$, where

$$R_\mu(\ell)(\ell') = \int_{\mathcal{X}} \left(\langle \ell, u \rangle - a_\mu(\ell)\right) \left(\langle \ell', u \rangle - a_\mu(\ell')\right) \mathrm{d}\mu(u) \qquad (\ell, \ell' \in \mathcal{X}^*).$$

If $\mathcal{X}$ is a function space, we call U Gaussian random field.

This definition does not immediately lead to a construction of a Gaussian measure on a general separable Banach space. There are two cases, in which we have techniques to construct a Gaussian measure on $\mathcal{X}$; those are $\mathbb{R}^k$ and separable Hilbert spaces. In finite dimensions, one can define a Gaussian measure in terms of a probability density function with respect to the product of a Lebesgue measure and a Dirac measure. On a separable Hilbert space, we can construct a series expansion, the so-called Karhunen-Loève expansion.

Definition 8.1.2. *Let $\mathcal{X}$ be a separable Hilbert space and $C : \mathcal{X} \to \mathcal{X}$ be a compact, self adjoint linear operator. Moreover, let $(\lambda_i, \varphi_i)_{i\in\mathbb{N}} \in (\mathbb{R} \times \mathcal{X})^{\mathbb{N}}$ be the eigenpairs of C sorted decreasingly with respect to the absolute value of the eigenvalue and $(\varphi_i)_{i\in\mathbb{N}}$ is orthonormal. Then, we can represent*

$$Cx = \sum_{i=1}^{\infty} \lambda_i \langle x, \varphi_i \rangle_{\mathcal{X}} \varphi_i \qquad (x \in \mathcal{X}),$$

see also Theorem 2.2.4. C is a trace class operator, if $(\lambda_i)_{i\in\mathbb{N}} \in \ell^1$.

Proposition 8.1.3. *Let $\mathcal{X}$ be a separable Hilbert space and $C : \mathcal{X} \to \mathcal{X}$ be a linear operator that is self-adjoint, non-negative, and trace class. We denote the eigenpairs of C by $(\lambda_i, \varphi_i)_{i\in\mathbb{N}} \in (\mathbb{R} \times \mathcal{X})^{\mathbb{N}}$; the eigenvalues are sorted decreasingly and $(\varphi_i)_{i\in\mathbb{N}}$ is orthonormal. Finally, let $m \in \mathcal{X}$ and $\xi \sim \mathrm{N}(0, 1^2)^{\otimes\mathbb{N}}$. Then,*

$$U := m + \sum_{i=1}^{\infty} \sqrt{\lambda_i} \xi_i \varphi_i$$

is distributed according to a Gaussian measure with mean m and covariance operator C.

Proof. Let $k \in \mathbb{N}$ and $U_k := m + \sum_{i=1}^{k} \sqrt{\lambda_i} \xi_i \varphi_i$. Moreover, let $x \in \mathcal{X}$ and $x_i := \langle x, \varphi_i \rangle_{\mathcal{X}}$ for $i \in \mathbb{N}$. We first study the distribution of $\langle x, U \rangle_{\mathcal{X}}$.

$$\begin{aligned}
\langle x, U_k \rangle_{\mathcal{X}} &= \left\langle x, m + \sum_{i=1}^{k} \sqrt{\lambda_i} \xi_i \varphi_i \right\rangle_{\mathcal{X}} \\
&= \langle x, m \rangle_{\mathcal{X}} + \left\langle x, \sum_{i=1}^{k} \sqrt{\lambda_i} \xi_i \varphi_i \right\rangle_{\mathcal{X}} \\
&= \langle x, m \rangle_{\mathcal{X}} + \sum_{i=1}^{k} \sqrt{\lambda_i} \langle x, \varphi_i \rangle_{\mathcal{X}} \xi_i \\
&= \langle x, m \rangle_{\mathcal{X}} + \sum_{i=1}^{k} \underbrace{\sqrt{\lambda_i} x_i \xi_i}_{\sim \mathrm{N}(0, \lambda_i x_i^2)}
\end{aligned}$$

converges weakly to the Gaussian distribution $\mathrm{N}\left(\langle x, m \rangle_{\mathcal{X}}, \sum_{i=1}^{\infty} \lambda_i x_i^2\right)$ $(k \to \infty)$, if the sum $\sum_{i=1}^{\infty} \lambda_i x_i^2$ is finite. (This can be shown with the Fourier transform of Gaussian measures, as the $(\xi_i)_{i\in\mathbb{N}}$ are mutually independent). By assumption, we have $(\lambda_i)_{i\in\mathbb{N}} \in \ell^1$ and also $(x_i^2)_{i\in\mathbb{N}} \in \ell^1$, since $\sum_{i=1}^{\infty} x_i^2 = \|x\|_{\mathcal{X}}^2 < \infty$. Hence, also $\sum_{i=1}^{\infty} \lambda_i x_i^2 < \infty$.

Next, we show that U takes values in $\mathcal{X}$ with probability one, i.e. $\mathbb{P}(\|U\|_{\mathcal{X}} < \infty) = 1$. By Parseval's identity, we have

$$\|U\|_{\mathcal{X}} = \sum_{i=1}^{\infty} |\langle U, \varphi_i \rangle_{\mathcal{X}}|^2 = \sum_{i=1}^{\infty} \lambda_i \xi_i^2$$

which is almost surely finite by Theorem 1.1.4 of [11], as $(\lambda_i)_{i\in\mathbb{N}} \in \ell^1$.

Now, we look at mean and covariance of U. We have

$$\begin{aligned} a_\mu(x) = \int_{\mathcal{X}} \langle x, U\rangle \mathrm{d}\mathbb{P} &= \langle x, m\rangle_{\mathcal{X}} + \int_{\mathbb{R}^{\mathbb{N}}} \sum_{i=1}^{\infty} \sqrt{\lambda_i} x_i \xi_i \mathrm{dN}(0,1)^{\otimes\mathbb{N}}(\xi) \\ &= \langle x, m\rangle_{\mathcal{X}} + \sum_{i=1}^{\infty} \sqrt{\lambda_i} x_i \underbrace{\int_{\mathbb{R}^{\mathbb{N}}} \xi_i \mathrm{dN}(0,1)^{\otimes\mathbb{N}}(\xi)}_{=0} \\ &= \langle x, m\rangle_{\mathcal{X}}\,, \end{aligned}$$

where we used the Fubini-Tonelli theorem to switch infinite sum and integral: Note that

$$\sum_{i=1}^{\infty} \int_{\mathbb{R}^{\mathbb{N}}} x|\sqrt{\lambda_i} x_i \xi_i| \mathrm{dN}(0,1)^{\otimes\mathbb{N}}(\xi) = \sum_{i=1}^{\infty} \sqrt{\frac{2}{\pi}} \leqslant \sqrt{\frac{2}{\pi}} \|\sqrt{\lambda_i}\|_2 \||x_i|\|_2$$

by Cauchy-Schwarz. Moreover, the upper bound on the RHS is finite, since $(x_i)_{i\in\mathbb{N}}, (\lambda_i)_{i\in\mathbb{N}} \in \ell^2$. Hence, $a_\mu = m$. Furthermore, we have for $x' \in \mathcal{X}$:

$$\begin{aligned} R_\mu(x)(x') &= \int_{\mathcal{X}} \left(\langle u, x\rangle_{\mathcal{X}} - a_\mu(x)\right)\left(\langle u, x'\rangle_{\mathcal{X}} - a_\mu(x')\right) \mathrm{d}\mu(u) \\ &= \int_{\mathbb{R}^{\mathbb{N}}} \left\langle x, \sum_{i=1}^{\infty} \sqrt{\lambda_i}\xi_i\varphi_i \right\rangle_{\mathcal{X}} \left\langle \sum_{j=1}^{\infty} \sqrt{\lambda_j}\xi_j\varphi_j, x' \right\rangle_{\mathcal{X}} \mathrm{dN}(0,1)^{\otimes\mathbb{N}}(\xi) \\ &= \int_{\mathbb{R}^{\mathbb{N}}} \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} \sqrt{\lambda_i}\sqrt{\lambda_j} \langle x, \varphi_i\rangle_{\mathcal{X}} \xi_i\xi_j \langle \varphi_j, x'\rangle_{\mathcal{X}} \mathrm{dN}(0,1)^{\otimes\mathbb{N}}(\xi) \\ &= \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} \sqrt{\lambda_i}\sqrt{\lambda_j} \langle x, \varphi_i\rangle_{\mathcal{X}} \underbrace{\int_{\mathbb{R}^{\mathbb{N}}} \xi_i\xi_j \mathrm{dN}(0,1)^{\otimes\mathbb{N}}(\xi)}_{=\mathbf{1}_{\{j\}}(i)} \langle \varphi_j, x'\rangle_{\mathcal{X}} \\ &= \sum_{i=1}^{\infty} \lambda_i \langle x, \varphi_i\rangle_{\mathcal{X}} \langle \varphi_i, x'\rangle_{\mathcal{X}} = \langle x, Cx'\rangle_{\mathcal{X}}, \end{aligned}$$

where we could remove the sum over j above due to mutual independence of the ξ_i, ξ_j with $i \neq j$. We exchanged sums and integral again using the Fubini-Tonelli theorem:

$$\begin{aligned} &\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} \int_{\mathbb{R}^{\mathbb{N}}} |\sqrt{\lambda_i}\sqrt{\lambda_j} \langle x, \varphi_i\rangle_{\mathcal{X}} \xi_i\xi_j \langle \varphi_j, x'\rangle_{\mathcal{X}}| \mathrm{dN}(0,1)^{\otimes\mathbb{N}}(\xi) \\ &\quad = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} |\sqrt{\lambda_i}\sqrt{\lambda_j} \langle x, \varphi_i\rangle_{\mathcal{X}} \langle \varphi_j, x'\rangle_{\mathcal{X}}| \cdot \frac{2}{\pi} \\ &\quad = \sum_{i=1}^{\infty} |\sqrt{\lambda_i} \langle x, \varphi_i\rangle_{\mathcal{X}}| \sum_{j=1}^{\infty} |\sqrt{\lambda_j} \langle \varphi_j, x'\rangle_{\mathcal{X}}| \cdot \frac{2}{\pi}, \end{aligned}$$

which is again finite, as $(x_i)_{i\in\mathbb{N}}, (x_i')_{i\in\mathbb{N}}, (\lambda_i)_{i\in\mathbb{N}} \in \ell^2$. □

Definition 8.1.4. *The expansion*

$$m + \sum_{i=1}^{\infty} \sqrt{\lambda_i}\xi_i\varphi_i$$

in Proposition 8.1.3 is called Karhunen–Loève expansion (KLE). In the same proposition, we denote $\mu =: \mathrm{N}(m, C)$.

We can understand the KLE as the function space version of a principal component analysis. Indeed, random fields are often discretised by representing them as a KLE and truncating the expansion. We now study two examples of Gaussian random fields in $\mathcal{L}^2$.

Example 8.1.5 (Gaussian random fields in 2 dimensions)**.** Let $D = [0,1]^2$, $\mathcal{X} := \mathcal{L}^2(D, \mathcal{B}D, \lambda_2)$, $\ell > 0$, and $\sigma^2 \geqslant 0$. We define the exponential covariance function

$$c_{\exp}(x,y) := \sigma^2 \exp\left(-\frac{\|x-y\|_2}{\ell}\right) \qquad (x,y \in D)$$

and the Gaussian covariance function

$$c_{\mathrm{N}}(x,y) := \sigma^2 \exp\left(-\frac{\|x-y\|_2^2}{2\ell^2}\right) \qquad (x,y \in D).$$

The parameter ℓ is called correlation length, σ^2 is called pointwise variance. We can now define the associated covariance operators for $c \in \{c_{\exp}, c_{\mathrm{N}}\}$, by

$$C : \mathcal{X} \to \mathcal{X}, \varphi \mapsto \int_D \varphi(x) c(x,\cdot) \mathrm{d}\lambda_2(x).$$

Well-definedness of these covariance operators can be shown with Mercer's Theorem. In Figure 8.1, we show discretised samples of Gaussian random fields with both covariance functions and $\ell \in \{0.05, 0.1, 1\}$. The random fields have been discretised by a 100^2-dimensional piecewise-constant finite element approximation of the eigenpairs of the respective covariance operator.

8.2 Monte Carlo techniques

8.2.1 Standard Monte Carlo

Monte Carlo techniques aim at approximating integrals of the form

$$\overline{g} := \int_{\mathcal{X}} g \mathrm{d}\mu,$$

where μ is a probability distribution on $(\mathcal{X}, \mathcal{B}\mathcal{X})$ and $g : (\mathcal{X}, \mathcal{B}\mathcal{X}) \to (\mathbb{R}, \mathcal{B}\mathbb{R})$ is an integrable function. Standard Monte Carlo approaches this problem by generating independent samples $U_1, U_2, ... \sim \mu$ and computing the estimator

$$\widehat{g}_M := \frac{1}{M} \sum_{m=1}^{M} g(U_m),$$

for some $M \in \mathbb{N}$. Alternatively, we can understand Monte Carlo as a technique allowing us to approximate the probability measure μ by the probability measure

$$\widehat{\mu}_M := \frac{1}{M} \sum_{m=1}^{M} \delta(\cdot - U_m).$$

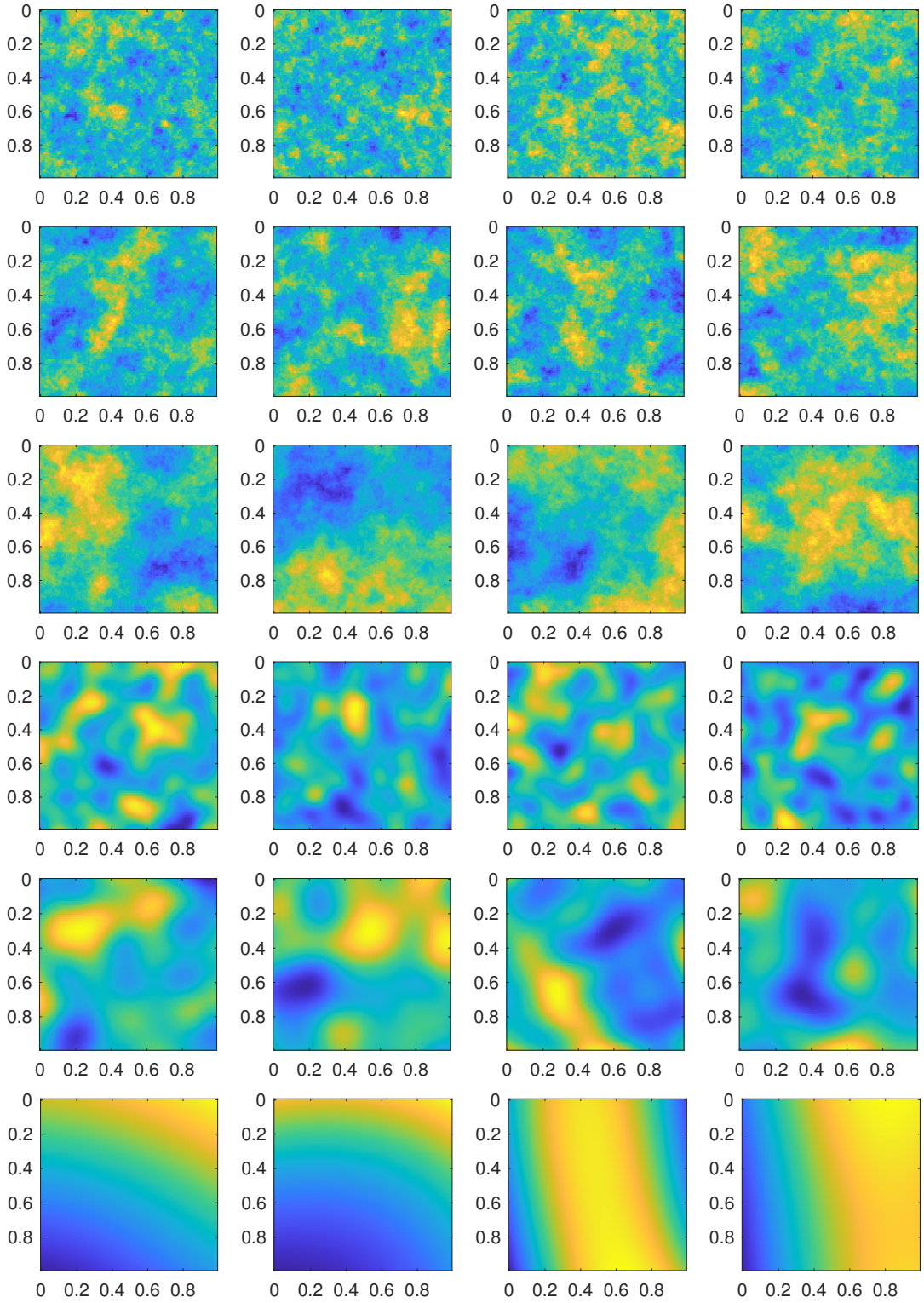


Figure 8.1: Each row represents four samples from the Gaussian random field with mean $m = 0$ and the following covariance operators (from top to bottom): exponential with $\ell = 0.05$, exponential with $\ell = 0.1$, exponential with $\ell = 1$, Gaussian with $\ell = 0.05$, Gaussian with $\ell = 0.1$, and Gaussian with $\ell = 1$.

The Monte Carlo estimator can be analysed using the (strong) law of large numbers. We know that

$$\widehat{g}_M \to \overline{g} \qquad (M \to \infty, \mathbb{P}\text{-a.s.}).$$

If in addition $\mathrm{Var}_\mu(g) := \int g^2 \mathrm{d}\mu - \left(\int g \mathrm{d}\mu\right)^2 < \infty$, we obtain the following convergence rate

$$\sqrt{\mathbb{E}\left[(\widehat{g}_M - \overline{g})^2\right]} = \frac{\sqrt{\mathrm{Var}_\mu(g)}}{\sqrt{M}}.$$

When thinking about standard algorithms for numerical quadrature (Gauss quadrature, Simpson's rule,...) the rate $O(M^{-1/2})$ appears to be quite slow. A composite Simpson's rule, e.g., for a very smooth function over $\mathcal{X} := [0,1]$ has an absolute error of $O(M^{-4})$. Its advantage over classical methods is that the rate is independent of the smoothness of the function and the dimension of its domain. Hence, Monte Carlo methods are especially useful in problems that are non-smooth and/or high-dimensional.

Unfortunately, standard Monte Carlo techniques are usually unsuitable for the approximation of posterior measures in Bayesian inverse problems: we are not able to sample independently from the posterior measure. Ideas:

- sample dependently from μ_{post} ($\to$ Markov chain Monte Carlo; this lecture) or
- sample independently from a different measure and correct by choosing unequal weights

  $$\widehat{g}_M := \sum_{m=1}^{M} w_m g(U_m),$$

  with $w_m \neq 1/M$, $m = 1, \ldots, M$ ($\to$ Importance Sampling; exercise sheet 4)

Markov chain Monte Carlo

In Markov chain Monte Carlo (MCMC), we generate a Markov chain $(U_m)_{m\in\mathbb{N}}$ that is stationary with respect to the posterior measure μ_{post} and Harris recurrent. In this case, we also have a law of large numbers

$$\frac{1}{M}\sum_{m=1}^{M} g(U_m) \to \int_{\mathcal{X}} g \mathrm{d}\mu_{\text{post}} \qquad (M \to \infty, \mathbb{P}\text{-a.s.}),$$

for some integrable $g : (\mathcal{X}, \mathcal{B}\mathcal{X}) \to (\mathbb{R}, \mathcal{B}\mathbb{R})$; see [30, Theorem 6.63]. We give a comparison of Monte Carlo and Markov chain Monte Carlo in Figure 8.2.

In the figure, we see that sampling a Markov chain can be less efficient than independent sampling – making MCMC not appearing very natural just yet. However, it is often easier to generate such a Markov chain than to sample independently from the posterior. In the following, we will first recap some definitions concerning Markov chains. Then, we will introduce the Metropolis–Hastings algorithm and show that is is stationary with respect to our measure of interest; say the posterior measure. We will not discuss ergodicity/Harris recurrence in this short introduction, but refer to the work by Robert and Casella [30].

Definition 8.2.1. *Let $(U_n)_{n=1}^{\infty}$ be a sequence of $\mathcal{X}$-valued random variables - so-called states. $(U_n)_{n=1}^{\infty}$ is called Markov chain, if for any $n \in \mathbb{N}$:*

$$\mathbb{P}(U_{n+1} \in \cdot | U_1 = u_1, U_2 = u_2, ..., U_{n-1} = u_{n-1}, U_n = u_n) = \mathbb{P}(U_{n+1} \in \cdot | U_n = u_n) \qquad (8.1)$$

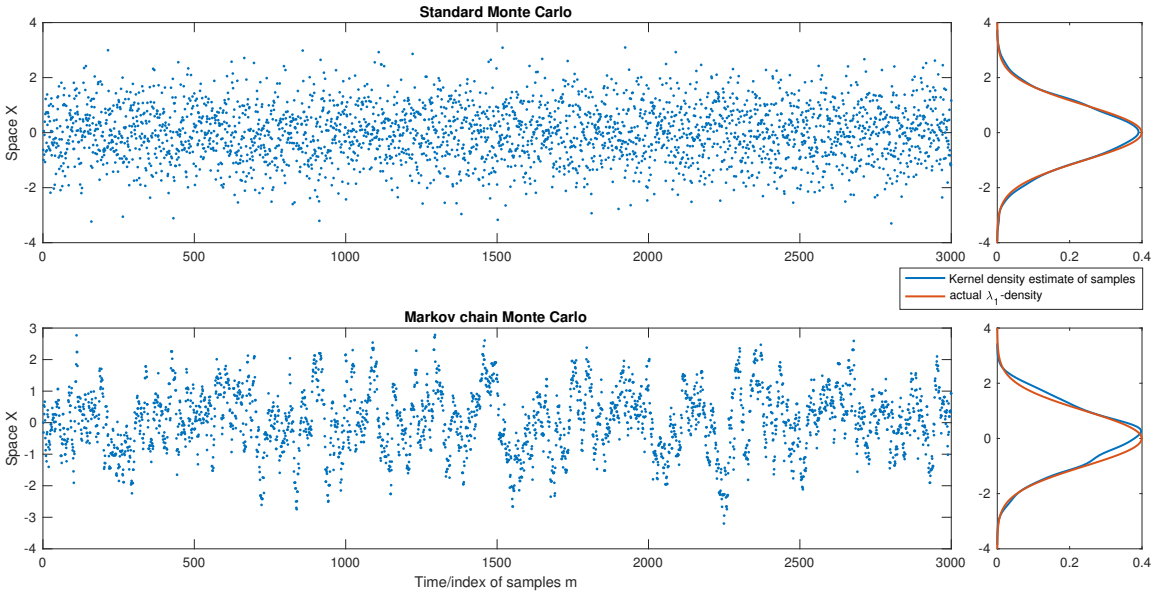


Figure 8.2: Comparison of Monte Carlo and Markov chain Monte Carlo samples. In the top row, we show 3000 independent samples of $\mathrm{N}(0, 1^2)$ and a kernel density estimate of these samples along with the true density. In the bottom row, we show 3000 samples generated with the Random Walk Metropolis algorithm targeting $\mathrm{N}(0, 1^2)$. The proposal kernel is $\mathrm{N}(\cdot, 0.5^2)$. The samples in the bottom row are clearly dependent.

for any $u_1, ..., u_{n-1} \in \mathcal{X}$. *A Markov chain is called* time-homogeneous, *if*

$$\mathbb{P}(U_2 \in \cdot | U_1 = u) = \mathbb{P}(U_{k+2} \in \cdot | U_{k+1} = u) \qquad (u \in \mathcal{X}, k \in \mathbb{N}). \tag{8.2}$$

and otherwise time-inhomogeneous. *A time-homogeneous Markov chain can be fully represented by a Markov kernel* $K : \mathcal{B}\mathcal{X} \times \mathcal{X} \to [0, 1]$*:*

$$K(B|u) = \mathbb{P}(U_{n+1} \in B | U_n = u) \qquad (B \in \mathcal{B}\mathcal{X}, u \in \mathcal{X}, n \in \mathbb{N}).$$

Let $\mu \in \mathrm{Prob}(\mathcal{X}, \mathcal{B}\mathcal{X})$ *be a probability measure. We denote the* composition of μ and K *by*

$$\mu K(B) := \int_{\mathcal{X}} K(B|u) \mathrm{d}\mu(u) \qquad (B \in \mathcal{B}\mathcal{X}).$$

The measure μ *is* stationary *w.r.t.* K*, if* $\mu K = \mu$*. Finally, we say, the Markov kernel* K *satisfies* detailed balance *w.r.t.* $\mu' \in \mathrm{Prob}(\mathcal{X}, \mathcal{B}\mathcal{X})$*, if*

$$\int_B K(A|u) \mathrm{d}\mu'(u) = \int_A K(B|u) \mathrm{d}\mu'(u) \qquad (A, B \in \mathcal{B}\mathcal{X}).$$

The detailed balance condition implies that the measure with respect to which it was shown is the stationary measure:

Lemma 8.2.2. *Let* $K : \mathcal{B}\mathcal{X} \times \mathcal{X} \to [0, 1]$ *be a Markov kernel that satisfies detailed balance with respect to* $\mu \in \mathrm{Prob}(\mathcal{X}, \mathcal{B}\mathcal{X})$*. Then,* K *is stationary w.r.t.* μ*.*

Proof. Exercise. □

We now define the Metropolis–Hastings Markov Kernel, discuss it, and show that it is stationary with respect to the target measure.

Definition 8.2.3 (Hastings 1970 [22])**.** *Let $\mu \in \mathrm{Prob}(\mathcal{X}, \mathcal{B}\mathcal{X})$ and ν be a σ-finite measure with $\mu \ll \nu$. Moreover let $g : (\mathcal{X}, \mathcal{B}\mathcal{X}) \to (\mathbb{R}, \mathcal{B}\mathbb{R})$ be a positive function with*

$$g = c \cdot \frac{\mathrm{d}\mu}{\mathrm{d}\nu},$$

for some $c \in (0, \infty)$. Moreover, let $Q : \mathcal{X} \times \mathcal{B}\mathcal{X} \to [0, 1]$ be a Markov kernel, given by a positive function $q : (\mathcal{X} \times \mathcal{X}, \mathcal{B}\mathcal{X} \otimes \mathcal{B}\mathcal{X}) \to (\mathbb{R}, \mathcal{B}\mathbb{R})$, with

$$Q(A|u) := \int_A q(u'|u)\mathrm{d}\nu(u') \qquad (A \in \mathcal{B}\mathcal{X}, u \in \mathcal{X}).$$

The Metropolis–Hastings Markov kernel is given by

$$K_{\mathrm{MH}}(A|u) := \delta(A-u) \int_{\mathcal{X}} (1-\alpha(u,u''))Q(\mathrm{d}u''|u) + \int_A \alpha(u,u')Q(\mathrm{d}u'|u) \qquad (u \in \mathcal{X}, A \in \mathcal{B}\mathcal{X}),$$

where

$$\alpha(u,u') = \min\left\{1, \frac{g(u')q(u|u')}{g(u)q(u'|u)}\right\}.$$

Interpreting this Markov kernel is rather difficult. Algorithmically, we can represent the Metropolis–Hastings MCMC method

1. Start with some initial value $U_1 \in \mathcal{X}$ (say a.s. constant); set $m \leftarrow 1$;
2. Sample $U^* \sim Q(\cdot|U_m)$; ('proposal step')
3. With probability $\alpha(U_m, U^*)$ set $U_{m+1} \leftarrow U^*$,
   otherwise $U_{m+1} \leftarrow U_{m+1}$; ('acceptance step')
4. Increment $m \leftarrow m + 1$ and go to 2.

When looking at K_{MH}, we see the proposal step in the Markov kernel Q and the acceptance step in the $(1 - \alpha)$ and the α.

Another remarkable observation is that we need to know the density g only up to a normalising constant. This is especially useful, when sampling from a posterior measure: we usually have only access to prior density and likelihood. Model evidence/normalising constant are not necessary.

Proposition 8.2.4. *K_{MH} satisfies detailed balance w.r.t. μ.*

Proof. Let $A, B \in \mathcal{B}\mathcal{X}$.

$$\begin{aligned}&\int_B K_{\mathrm{MH}}(A|u)\mathrm{d}\mu(u)\\ &\quad= \int_B \delta(A-u)\int_{\mathcal{X}}(1-\alpha(u,u''))Q(\mathrm{d}u''|u) + \int_A \alpha(u,u')Q(\mathrm{d}u'|u)\mathrm{d}\mu(u).\end{aligned}$$

We discuss the two parts of this sum one after another. We first have

$$\begin{aligned}
&\int_B \delta(A-u) \int_{\mathcal{X}} (1-\alpha(u,u''))Q(\mathrm{d}u''|u)\mathrm{d}\mu(u) \\
&\quad = \int_{\mathcal{X}} \mathbf{1}_{A\cap B}(u) \int_{\mathcal{X}} (1-\alpha(u,u''))Q(\mathrm{d}u''|u)g(u)\mathrm{d}\nu(u) \\
&\quad = \int_A \delta(B-u) \int_{\mathcal{X}} (1-\alpha(u,u''))Q(\mathrm{d}u''|u)\mathrm{d}\mu(u).
\end{aligned}$$

Secondly,

$$\begin{aligned}
&\int_B \int_A \alpha(u,u')Q(\mathrm{d}u'|u)\mathrm{d}\mu(u) \\
&\quad = \int_B \int_A \min\left\{1, \frac{g(u')q(u|u')}{g(u)q(u'|u)}\right\} q(u'|u)\mathrm{d}\nu(u')\frac{g(u)}{c}\mathrm{d}\nu(u) \\
&\quad = \int_B \int_A \min\left\{g(u)q(u'|u), g(u')q(u|u')\right\} \mathrm{d}\nu(u')\frac{1}{c}\mathrm{d}\nu(u) \\
&\quad = \int_A \int_B \min\left\{1, \frac{g(u)q(u'|u)}{g(u')q(u|u')}\right\} \frac{g(u')}{c} q(u|u')\mathrm{d}\nu(u)\mathrm{d}\nu(u') \\
&\quad = \int_A \int_B \alpha(u',u)Q(\mathrm{d}u|u')\mathrm{d}\mu(u').
\end{aligned}$$

Combining these two results gives us detailed balance. □

We finish by giving typical examples for proposal kernels Q used in Metropolis–Hastings MCMC.

Example 8.2.5 (Independence Sampler)**.** Let $\rho \in \mathrm{Prob}(\mathcal{X}, \mathcal{B}\mathcal{X})$. The Metropolis-Hastings algorithm with proposal kernel

$$Q(\cdot|u) = \rho \qquad (u \in \mathcal{X})$$

is called independence sampler. The acceptance probability is given by

$$\alpha(u,u') = \min\left\{1, \frac{g(u')q(u)}{g(u)q(u')}\right\},$$

where $q = \mathrm{d}\rho/\mathrm{d}\nu$.

In a Bayesian inverse problem with prior $\mu_0 \in \mathrm{Prob}(\mathcal{X}, \mathcal{B}\mathcal{X})$ and likelihood $L(f_n|\cdot)$, we can choose $\rho := \nu := \mu_0$. In this case, the acceptance probability simplifies to

$$\alpha(u,u') = \min\left\{1, \frac{L(f_n|u')}{L(f_n|u)}\right\}.$$

Please note that the independence sampler proposes moves independently of the current position. This does not imply that the generated samples are independent. The acceptance step couples the samples.

Example 8.2.6 (Random Walk; Metropolis et al. 1953 [27])**.** Let $\rho \in \mathrm{Prob}(\mathcal{X}, \mathcal{B}\mathcal{X})$ have a symmetric density $q' = \mathrm{d}\rho/\mathrm{d}\nu$, i.e. $q' = q'(-\cdot)$. The Metropolis-Hastings algorithm with proposal kernel

$$Q(\cdot|u) = \rho(\cdot - u) \qquad (u \in \mathcal{X})$$

is called Random Walk Metropolis sampler. The acceptance probability is given by

$$\alpha(u, u') = \min\left\{1, \frac{g(u')}{g(u)}\right\}.$$

Note that the acceptance probability is independent of the proposal distribution; indeed, it cancels: $q(u|u') = q'(u - u') = q'(u' - u) = q(u'|u)$.

Example 8.2.7 (Preconditioned Crank–Nicolson MCMC; Cotter et al. 2013 [16])**.** Let $\mathcal{X}$ be a separable Hilbert space and let $\mu_0 = \mathrm{N}(0, \mathcal{C}) \in \mathrm{Prob}(\mathcal{X}, \mathcal{B}\mathcal{X})$ for some suitable operator $\mathcal{C} : \mathcal{X} \to \mathcal{X}$. We consider the (BIP) with prior μ_0 and likelihood $L(f_n|\cdot)$. Let $\beta \in (0, 1)$ The Metropolis-Hastings algorithm with proposal kernel

$$Q(\cdot|u) := \mathrm{N}(\sqrt{1 - \beta^2}u, \beta^2\mathcal{C})$$

is called preconditioned Crank–Nicolson algorithm (pCN-MCMC). The acceptance probability is given by

$$\alpha(u, u') = \min\left\{1, \frac{L(f_n|u')}{L(f_n|u)}\right\}.$$

This method is particularly useful in high- and infinite dimension, where the random walk algorithm cannot be applied. Proving that α is the correct acceptance probability is rather simple in finite dimensions, not quite as easy in infinite dimensions.

The method is referred to as pCN MCMC as the proposal can be derived as a Crank–Nicolson discretisation of some S(P)DE.

Bibliography

[1] Y. A. Abramovich and C. D. Aliprantis, *An Invitation to Operator Theory*, Graduate Studies in Mathematics, American Mathematical Society, 2002.

[2] R. A. Adams and J. J. F. Fournier, *Sobolev Spaces*, Elsevier Science, Singapore, 2003.

[3] S. Agapiou, O. Papaspiliopoulos, D. Sanz-Alonso, and A. M. Stuart, *Importance sampling: intrinsic dimension and computational cost*, Statist. Sci., 32 (2017), pp. 405–431.

[4] C. D. Aliprantis and K. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, Springer, 2006.

[5] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of Bounded Variation and Free Discontinuity Problems*, Clarendon Press, 2000.

[6] R. B. Ash and C. A. Doléans-Dade, *Probability & Measure Theory*, Harcourt Academic Press, 2000.

[7] A. B. Bakushinskii, *Remarks on the choice of regularization parameter from quasioptimality and relation tests*, Zhurnal Vychislitel'noï Matematiki i Matematicheskoï Fiziki, 24 (1984), pp. 1258–1259.

[8] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, 2011.

[9] M. Benning and M. Burger, *Modern regularization methods for inverse problems*, Acta Numerica, 27 (2018), pp. 1–111.

[10] P. Billingsley, *Probability and Measure*, John Wiley and Sons, second ed., 1986.

[11] V. I. Bogachev, *Gaussian measures*, vol. 62 of Mathematical Surveys and Monographs, American Mathematical Society, Providence, RI, 1998.

[12] B. Bollobás, *Linear Analysis: An Introductory Course*, Cambridge University Press, Cambridge, second ed., 1999.

[13] K. Bredies and D. A. Lorenz, *Mathematical Image Processing*, Springer, 2018.

[14] M. Burger and S. Osher, *Convergence rates of convex variational regularization*, Inverse Problems, 20 (2004), p. 1411.

[15] ——, *A guide to the tv zoo*, in Level-Set and PDE-based Reconstruction Methods, M. Burger and S. Osher, eds., Springer, 2013.

[16] S. L. Cotter, G. O. Roberts, A. M. Stuart, and D. White, *MCMC Methods for Functions: Modifying Old Algorithms to Make Them Faster*, Statist. Sci., 28 (2013), pp. 424–446.

[17] R. T. Cox, *Probability, frequency and reasonable expectation*, American Journal of Physics, 14 (1946), pp. 1–13.

[18] N. Dunford and J. T. Schwartz, *Linear Operators, Part 1: General Theory*, Wiley Interscience Publishers, 1988.

[19] I. Ekeland and R. Témam, *Convex Analysis and Variational Problems*, 1976.

[20] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of inverse problems*, vol. 375, Springer Science & Business Media, 1996.

[21] C. W. Groetsch, *Stable approximate evaluation of unbounded operators*, Springer, 2006.

[22] W. K. Hastings, *Monte Carlo Sampling Methods Using Markov Chains and Their Applications*, Biometrika, 57 (1970), pp. 97–109.

[23] J. Hunter and B. Nachtergaele, *Applied Analysis*, World Scientific Publishing Company Incorporated, 2001.

[24] J. A. Iglesias, G. Mercier, and O. Scherzer, *A note on convergence of solutions of total variation regularized linear inverse problems*, Inverse Problems, 34 (2018), p. 055011.

[25] A. Klenke, *Probability Theory: A comprehensive Course*, Springer, 2014.

[26] J. Leao, D, M. Fragoso, and P. Ruffino, *Regular conditional probability, disintegration of probability and Radon spaces*, Proyecciones, 23 (2004), pp. 15–29.

[27] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, *Equation of State Calculations by Fast Computing Machines*, J. Chem. Phys., 21 (1953), pp. 1087–1092.

[28] A. W. Naylor and G. R. Sell, *Linear Operator Theory in Engineering and Science*, Springer Science & Business Media, 2000.

[29] Y. V. Prokhorov, *Convergence of random processes and limit theorems in probability theory*, Theory of Probability & Its Applications, 1 (1956), pp. 157–214.

[30] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*, Springer, 2004.

[31] L. I. Rudin, S. Osher, and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Physica D: Nonlinear Phenomena, 60 (1992), pp. 259–268.

[32] W. Rudin, *Functional Analysis*, International series in pure and applied mathematics, McGraw-Hill, 1991.

[33] K. Saxe, *Beginning Functional Analysis*, Springer, 2002.

[34] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen, *Variational Methods in Imaging*, Springer, 2009.

[35] A. M. Stuart, *Inverse problems: a Bayesian perspective*, Acta Numerica, 19 (2010), pp. 451–559.

[36] T. Tao, *Epsilon of Room, One*, vol. 1, American Mathematical Soc., 2010.

[37] E. Zeidler, *Applied Functional Analysis: Applications to Mathematical Physics*, vol. 108 of Applied Mathematical Sciences Series, Springer, 1995.

[38] ——, *Applied Functional Analysis: Main Principles and Their Applications*, vol. 109 of Applied Mathematical Sciences Series, Springer, 1995.