Mathematical Tripos Part IB: Easter 2001

Numerical Analysis – Lecture 7[1]

5 Linear least squares

5.1 Statement of the problem

Consider the problem of finding a vector $\boldsymbol{x} \in \mathbb{R}^n$ such that $A\boldsymbol{x} = \boldsymbol{b}$ where a matrix $A \in \mathbb{R}^{m\times n}$ and a vector $\boldsymbol{b} \in \mathbb{R}^m$ are given and $m > n$. When there are more equations than unknowns, the system $A\boldsymbol{x} = \boldsymbol{b}$ is called *overdetermined*. In general, an overdetermined system has no solution, but we would like to have $A\boldsymbol{x}$ and $\boldsymbol{b}$ close in a sense. Problems of this form occur frequently when we collect m observations (which, typically, are prone to measurement error) and wish to exploit them to form an n-variable linear model (e.g., trying to put some planet observartions on an ellipse). Choosing the Euclidean distance as such a proximity measure, one obtains

The *least squares problem*. Given $A \in \mathbb{R}^{m\times n}$ and $\boldsymbol{b} \in \mathbb{R}^m$, find $\boldsymbol{x}^* \in \mathbb{R}^n$ which minimizes $\|A\boldsymbol{x} - \boldsymbol{b}\|$.

Let $A = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n]$ with $\boldsymbol{a}_i \in \mathbb{R}^m$. Then the vectors $A\boldsymbol{x} = \sum_{i=1}^n x_i \boldsymbol{a}_i$, where $\boldsymbol{x}$ runs through $\mathbb{R}^n$, form an n-dimensional plane in $\mathbb{R}^m$, say $\mathcal{A}$, spanned by $(\boldsymbol{a}_i)_{i=1}^n$. Given $\boldsymbol{b} \in \mathbb{R}^m$, the minimal distance $\|\boldsymbol{a}^* - \boldsymbol{b}\|$ between $\boldsymbol{b}$ and vectors in $\mathcal{A}$ is attained when $\boldsymbol{a}^*$ is the foot of the perpendicular from $\boldsymbol{b}$ onto the plane $\mathcal{A}$.

Theorem. Given $\boldsymbol{b} \in \mathbb{R}^m$ and $\mathcal{A} = \mathrm{span}(\boldsymbol{a}_i)_{i=1}^n$, the vector $\boldsymbol{a}^*$ minimizes $\|\boldsymbol{a} - \boldsymbol{b}\|$ over all $\boldsymbol{a} \in \mathcal{A}$ iff $\langle \boldsymbol{a}^* - \boldsymbol{b}, \boldsymbol{a}\rangle = 0$ for all $\boldsymbol{a} \in \mathcal{A}$.

Proof. Given $\boldsymbol{a}_*, \boldsymbol{a}, \boldsymbol{b}$, consider the quadratic polynomial in λ,

$$p(\lambda) = \|(\boldsymbol{a}^* - \lambda\boldsymbol{a}) - \boldsymbol{b}\|^2 = \|\boldsymbol{a}^* - \boldsymbol{b}\|^2 - 2\lambda\langle \boldsymbol{a}^* - \boldsymbol{b}, \boldsymbol{a}\rangle + \lambda^2\|\boldsymbol{a}\|^2.$$

Its least value occurs when $p'(\lambda_*) = 0$, so that

$$\lambda_* = \frac{\langle \boldsymbol{a}_* - \boldsymbol{b}, \boldsymbol{a}\rangle}{\|\boldsymbol{a}\|^2}, \quad p(\lambda_*) = \|\boldsymbol{a}^* - \boldsymbol{b}\|^2 - \lambda_*^2\|\boldsymbol{a}\|^2.$$

If $\langle \boldsymbol{a}^* - \boldsymbol{b}, \boldsymbol{a}\rangle$ is nonzero for some $\boldsymbol{a}$, so is λ_*, and then $p(\lambda_*) < p(0) = \|\boldsymbol{a}^* - \boldsymbol{b}\|^2$, i.e., $\boldsymbol{a}^*$ is not optimal. Conversely, if $\langle \boldsymbol{a}^* - \boldsymbol{b}, \boldsymbol{a}\rangle = 0$ for all $\boldsymbol{a} \in \mathcal{A}$, then, for any $\boldsymbol{y} \in \mathcal{A}$, with $\boldsymbol{a} := \boldsymbol{a}^* - \boldsymbol{y}$, we find that

$$\|\boldsymbol{y} - \boldsymbol{b}\|^2 = \|(\boldsymbol{a}^* - \boldsymbol{a}) - \boldsymbol{b}\|^2 = \|\boldsymbol{a}^* - \boldsymbol{b}\|^2 + \|\boldsymbol{a}\|^2 > \|\boldsymbol{a}^* - \boldsymbol{b}\|^2,$$

unless $\boldsymbol{y} = \boldsymbol{a}^*$. $\square$

Corollary. $\boldsymbol{x}^* \in \mathbb{R}^n$ minimizes $\|A\boldsymbol{x} - \boldsymbol{b}\|$ iff $\langle A\boldsymbol{x}^* - \boldsymbol{b}, A\boldsymbol{x}\rangle = 0$ for all $\boldsymbol{x} \in \mathbb{R}^n$.

Corollary. $\boldsymbol{x}^* \in \mathbb{R}^n$ minimizes $\|A\boldsymbol{x} - \boldsymbol{b}\|$ iff $A^{\mathrm{T}}A\boldsymbol{x}^* = A^{\mathrm{T}}\boldsymbol{b}$.

The equation $A^{\mathrm{T}}A\boldsymbol{x} = A^{\mathrm{T}}\boldsymbol{b}$ is called the *normal equation*.

5.2 Normal equations

One way of finding optimal $\boldsymbol{x}$ is by solving the $n \times n$ linear system $A^{\mathrm{T}}A\boldsymbol{x} = A^{\mathrm{T}}\boldsymbol{b}$ – the method of *normal equations*. This approach is popular in many applications. However, there are three disadvantages. Firstly, $A^{\mathrm{T}}A$ might be singular, secondly sparse A might be replaced by a dense $A^{\mathrm{T}}A$ and, finally, forming $A^{\mathrm{T}}A$ might lead to loss of accuracy. Thus, suppose that our computer works in the IEEE arithmetic standard (≈ 15 significant digits) and let

$$A = \begin{bmatrix} 10^8 & -10^8 \\ 1 & 1 \end{bmatrix} \qquad \Longrightarrow \qquad A^{\mathrm{T}}A = \begin{bmatrix} 10^{16}+1 & -10^{16}+1 \\ -10^{16}+1 & 10^{16}+1 \end{bmatrix} \approx 10^{16}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

Given $\boldsymbol{b} = [0, 2]^{\mathrm{T}}$ the solution of $A\boldsymbol{x} = \boldsymbol{b}$ is $[1, 1]^{\mathrm{T}}$, as can be easily found by Gaussian elimination. However, our computer 'believes' that $A^{\mathrm{T}}A$ is singular!

[1]Corrections and suggestions to these notes should be emailed to `a.shadrin@damtp.cam.ac.uk`. All handouts are available on the WWW at the URL `http://www.damtp.cam.ac.uk/user/na/PartIB/Handouts.html`.

5.3 QR and least squares

Lemma Let A be any $m \times n$ matrix and let $\boldsymbol{b} \in \mathbb{R}^m$. The vector $\boldsymbol{x} \in \mathbb{R}^n$ minimises $\|A\boldsymbol{x} - \boldsymbol{b}\|$ iff it minimises $\|\Omega A\boldsymbol{x} - \Omega\boldsymbol{b}\|$ for an arbitrary $m \times m$ orthogonal matrix Ω.

Proof. Given an arbitrary vector $\boldsymbol{v} \in \mathbb{R}^m$, we have

$$\|\Omega\boldsymbol{v}\|^2 = \boldsymbol{v}^{\mathrm{T}}\Omega^{\mathrm{T}}\Omega\boldsymbol{v} = \boldsymbol{v}^{\mathrm{T}}\boldsymbol{v} = \|\boldsymbol{v}\|^2.$$

In particular, $\|\Omega A\boldsymbol{x} - \Omega\boldsymbol{b}\| = \|A\boldsymbol{x} - \boldsymbol{b}\|$. □

An irrelevant, yet important remark The property that orthogonal matrices leave the Euclidean distance intact is called *isometry* and it has many important ramifications throughout mathematics and mathematical physics.

Method of solution Suppose that $A = QR$, with an orthogonal Q and an upper triangular R. Because of the lemma, letting $\Omega := Q^{\mathrm{T}}$,

$$\|A\boldsymbol{x} - \boldsymbol{b}\| = \|Q^{\mathrm{T}}(A\boldsymbol{x} - \boldsymbol{b})\| = \|R\boldsymbol{x} - Q^{\mathrm{T}}\boldsymbol{b}\|,$$

therefore we seek $\boldsymbol{x} \in \mathbb{R}^n$ that minimises $\|R\boldsymbol{x} - Q^{\mathrm{T}}\boldsymbol{b}\|$.
Suppose for simplicity that $\operatorname{rank} R = \operatorname{rank} A = n$. Then the bottom $m - n$ rows of R are zero. Therefore we find $\boldsymbol{x}$ by solving the (nonsingular) linear system given by the first n equations of $R\boldsymbol{x} = Q^{\mathrm{T}}\boldsymbol{b}$. Similar (although more complicated) algorithm applies when $\operatorname{rank} R \leq n - 1$. Note, recalling our former remark, that we don't require Q explicitly and need to evaluate only $Q^{\mathrm{T}}\boldsymbol{b}$.

6 Polynomial interpolation

6.1 The interpolation problem

Given $n + 1$ distinct real points $x_0, x_1, \ldots, x_n$ and real numbers $f_0, f_1, \ldots, f_n$, we seek a function $p : \mathbb{R} \to \mathbb{R}$ such that $p(x_i) = f_i$, $i = 0, 1, \ldots, n$. Such a function is called an *interpolant*.
We denote by $\mathbb{P}_n[x]$ the set of all real polynomials of degree at most n and observe that each $p \in \mathbb{P}_n[x]$ is uniquely defined by its $n + 1$ coefficients. In other words, we have $n + 1$ degrees of freedom, while interpolation at $x_0, x_1, \ldots, x_n$ constitutes $n + 1$ conditions. This, intuitively, justifies seeking an interpolant from $\mathbb{P}_n[x]$.

6.2 The Lagrange formula

Although, in principle, we may solve a linear problem with $n + 1$ unknowns to determine a polynomial interpolant, this can be accomplished more easily by using the explicit *Lagrange formula.* We claim that

$$p(x) = \sum_{k=0}^{n} f_k \prod_{\substack{\ell=0 \\ \ell \neq k}}^{n} \frac{x - x_\ell}{x_k - x_\ell}, \qquad x \in \mathbb{R}.$$

Note that $p \in \mathbb{P}_n[x]$, as required. We wish to show that it interpolates the data. Define

$$L_k(x) := \prod_{\substack{\ell=0 \\ \ell \neq k}}^{n} \frac{x - x_\ell}{x_k - x_\ell}, \qquad k = 0, 1, \ldots, n$$

(*Lagrange cardinal polynomials*). It is trivial to verify that $L_k(x_k) = 1$ and $L_k(x_j) = 0$ for $k \neq j$, hence

$$p(x_j) = \sum_{k=0}^{n} f_k L_k(x_j) = f_j, \qquad j = 0, 1, \ldots, n,$$

and p is an interpolant,

Uniqueness. Suppose that both $p \in \mathbb{P}_n[x]$ and $q \in \mathbb{P}_n[x]$ interpolate to the same $n + 1$ data. Then the nth degree polynomial $p - q$ vanishes at $n + 1$ distinct points. But the only nth-degree polynomial with $\geq n + 1$ zeros is the zero polynomial. Therefore $p - q \equiv 0$ and the interpolating polynomial is unique.