Mathematical Tripos Part IB: Easter Term 2002

Numerical Analysis – Lecture 7

7 Linear least squares

7.1 Statement of the problem

Consider the problem of finding a vector $x \in \mathbb{R}^n$ such that $Ax = b$ where a matrix $A \in \mathbb{R}^{m\times n}$ and a vector $b \in \mathbb{R}^m$ are given and $m > n$. When there are more equations than unknowns, the system $Ax = b$ is called *overdetermined*. In general, an overdetermined system has no solution, but we would like to have Ax and 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 observations on an ellipse). Choosing the Euclidean distance as a measure of closeness, one obtains

Problem 7.1 (Least squares in $\mathbb{R}^m$) Given $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$, find

$$x^* = \arg\min_{x\in\mathbb{R}^n} \|Ax - b\|,$$

i.e., find (the argument) $x^* \in \mathbb{R}^n$ which minimizes (the functional) $\|Ax - b\|$.

Discussion 7.2 Let $A = [a_1, \dots, a_n]$ with $a_i \in \mathbb{R}^m$. Then the vectors $Ax = \sum_{j=1}^n x_j a_j$, where x runs through $\mathbb{R}^n$, form an n-dimensional plane in $\mathbb{R}^m$, say $\mathcal{A}$, spanned by $(a_j)_{j=1}^n$. Problem 7.1 is the problem of finding

$$a^* = \arg\min_{a\in\mathcal{A}} \|a - b\|.$$

This problem has a clear geometric solution: given $b \in \mathbb{R}^m$, the minimal distance $\|a^* - b\|$ between b and vectors in $\mathcal{A}$ is attained when a^* is the foot of the perpendicular from b onto the plane $\mathcal{A}$, i.e. when

$$a^* - b \perp \mathcal{A}, \qquad \mathcal{A} = \operatorname{span}(a_i)$$

The next theorem puts these arguments into precise analytic form. Since the proof uses nothing more than the properties of the scalar product in $\mathbb{R}^m$, it costs nothing to replace $\mathbb{R}^m$ by a general inner product space $\mathbb{X}$.

Problem 7.3 (Least squares in $\mathbb{X}$) Given $b \in \mathbb{X}$ and $(a_j)_{j=1}^n \in \mathbb{X}$, find

$$x^* = \arg\min_{x\in\mathbb{R}^n} \| \textstyle\sum_{j=1}^n x_j a_j - b\|.$$

The element $a^* = \sum_{j=1}^n x_j^* a_j$ is called the best least squares approximation to b from $\mathcal{A} := \operatorname{span}(a_j)$.

Theorem 7.4 *Let $\mathbb{X}$ be an inner product space, and let $b \in \mathbb{X}$ and a subspace $\mathcal{A} \subset \mathbb{X}$ be given. Then*

$$a^* = \arg\min_{a\in\mathcal{A}} \|a - b\| \quad \Leftrightarrow \quad (a^* - b, a) = 0 \quad \forall a \in \mathcal{A}. \tag{7.1}$$

Proof. Given $a^*, a, b \in \mathbb{X}$, consider the quadratic polynomial in λ,

$$p(\lambda) := \|(a^* - \lambda a) - b\|^2 = \|a^* - b\|^2 - 2\lambda(a^* - b, a) + \lambda^2\|a\|^2.$$

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

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

If $(a^* - b, a) \neq 0$ for some $a \in \mathcal{A}$, so is λ_*, and then $p(\lambda_*) < p(0) = \|a^* - b\|^2$, i.e., a^* is not optimal.

Conversely, if $(a^* - b, a) = 0$ for all $a \in \mathcal{A}$, then, for any $a' \in \mathcal{A}$, with $a := a^* - a'$, we find that

$$\|a' - b\|^2 := \|(a^* - a) - b\|^2 = \|(a^* - b) - a\|^2 = \|a^* - b\|^2 + \|a\|^2 > \|a^* - b\|^2,$$

i.e., a^* is the best approximant. ∎

Corollary 7.5 *If a^* is the best least squares approximation to $b \in \mathbb{X}$ from $\mathcal{A}$, then*

$$\|b - a^*\|^2 + \|a^*\|^2 = \|b\|^2, \quad \textit{in particular} \quad \|a^*\| \leq \|b\|,$$

the latter inequality being strict for $b \in \mathbb{X} \setminus \mathcal{A}$.

7.2 Normal equations

If $\mathcal{A} = \operatorname{span}(a_i)$, then writing $a^* = \sum_j x_j^* a_j$ and running a in (7.1) through the basis finctions a_i we obtain a linear system of equations for determining the coefficients x^*

$$\textstyle\sum_{j=1}^n (a_i, a_j) x_j^* = (a_i, b), \quad i = \overline{1, n},$$

so that we may reformulate Theorem 7.4 in the following way.

Theorem 7.6 *Let $(a_i)_{i=1}^n \subset \mathbb{X}$ and let $b \in \mathbb{X}$. Then*

$$x^* = \arg\min_{x \in \mathbb{R}^n} \| \textstyle\sum_j x_j a_j - b\| \quad \Leftrightarrow \quad Gx^* = y,$$

where

$$G := [(a_i, a_j)]_{i,j=1}^n \in \mathbb{R}^{n \times n}, \qquad y := [(a_i, b)] \in \mathbb{R}^n.$$

The system $Gx = y$ is called the *normal* equations, G is the *Gram* matrix, x^* is the *normal* solution.

Coming back to our particular problem of least squares solution to $Ax = b$ we obtain

Corollary 7.7 *Let $b \in \mathbb{R}^m$ and $A \in \mathbb{R}^{m \times n}$. Then*

$$x^* = \arg\min_{x \in \mathbb{R}^n} \|Ax - b\| \quad \Leftrightarrow \quad Ax - b \perp a_i \quad \Leftrightarrow \quad A^T A x^* = A^T b.$$

However, the method of normal equations $A^T A x = A^T b$ for finding an optimal x^* has several disadvantages. The main one is ill-conditioning.

Example 7.8 (Polynomial least squares) Given $f \in C[0, 1]$, find an algebraic polynomial $p^* \in \mathcal{P}_n$ of degree n such that

$$p^* = \arg\min_{p \in \mathcal{P}_n} \sum_{i=0}^m |p(t_i) - f(t_i)|^2, \quad t_i = i/m.$$

If one seeks $p(t) = \sum_{j=0}^n \gamma_j t^j$, then $A = [(i/m)^j]$ and $A^T A$ is very close to the so-called Hilbert matrix, the monster that explodes all the algorithms.

7.3 QR and least squares

An alternative is provided by the QR factorization. Suppose that $A = QR$, with an orthogonal $Q \in \mathbb{R}^{m \times m}$ and an upper triangular $R \in \mathbb{R}^{m \times n}$. Then, since orthogonal mapping is length-preserving, i.e., $\|Q^T v\|^2 = \|v\|^2$ for any $v \in \mathbb{R}^m$, we have

$$\|Ax - b\| = \|Q^T (Ax - b)\| = \|Rx - Q^T b\|$$

therefore we may seek $x \in \mathbb{R}^n$ that minimises $\|Rx - Q^T b\|$.

Suppose for simplicity that $\operatorname{rank} R = \operatorname{rank} A = n$. Then the bottom $m - n$ rows of R are zero. Therefore we find x by solving the (nonsingular) linear system given by the first n equations of

$$Rx = Q^T b.$$

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 $Q^T b$.