MA5233 Computational Mathematics

Lecture 9: Least Squares

Simon Etter

2019/2020

Least Squares

Least squares problem

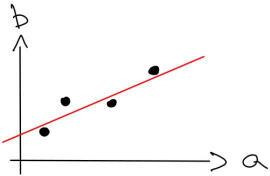
Given $A \in \mathbb{K}^{m \times n}$ and $b \in \mathbb{K}^m$ with $m > n$, find

$$x := \arg\min \|Ax - b\|_2.$$

Example

Given data vectors $a, b \in \mathbb{R}^m$ defining points (a_i, b_i) in the 2d plane, find $c_0, c_1 \in \mathbb{R}$ minimising

$$\sum_{i=1}^{m} (c_0 + c_1\, a_i - b_i)^2.$$



This is a least squares problem with

$$A = \begin{pmatrix} 1 & a \end{pmatrix}, \qquad b = b, \qquad x = \begin{pmatrix} c_0 & c_1 \end{pmatrix}^T.$$

Least Squares

Solution via QR factorisation

Let $QR := A$ with

$$Q = \begin{pmatrix} Q_1 & Q_2 \end{pmatrix} \in \mathbb{K}^{m \times m} \text{ orthogonal,}$$
$$R = \begin{pmatrix} R_1 \\ 0 \end{pmatrix} \in \mathbb{K}^{m \times n} \text{ upper-triangular.}$$

Since Q is orthogonal, we have

$$\|Ax - b\|_2 = \left\|Q\left(Rx - Q^H b\right)\right\|_2 = \|Rx - Q^H b\|_2,$$

or in block-matrix form

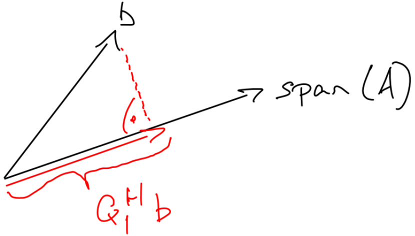
$$\left\|\begin{pmatrix} R_1 \\ 0 \end{pmatrix} x - \begin{pmatrix} Q_1^H b_1 \\ Q_2^H b_2 \end{pmatrix}\right\|_2 .$$

Clearly, solution to least squares problem is given by $\boxed{x = R_1^{-1} Q_1^H b.}$
Note: $A = Q_1 R_1$ is thin QR factorisation of A.

Least Squares

Interpretation of $x = R^{-1} Q_1^H b$

- $p := Q_1^H b$: project b onto span(A).
- $x = R_1^{-1} p$: solve for coordinates of p with respect to A.



Solving least squares problem in Julia

See `9_least_squares.jl`.

Least Squares

Conditioning of least squares problem

Two parameters:

- $\kappa(A) := \frac{\sigma_{\max}}{\sigma_{\min}}$ (measures linear independence of columns of A)
- $\theta := \cos^{-1}\left(\frac{\|Ax\|}{\|b\|}\right)$ (measures degree of orthogonality of b to span(A))

Literature provides several distinct but related conditioning estimates.

General theme is that problem is well-conditioned if $\kappa(A)$ and θ are small.

See references at the end of these slides for more detail.

Stability of solving least squares via QR factorisation

Backward stable.

Least Squares

Real-world example: pricing of diamonds

Price v of diamond is primarily determined by four parameters:

- $w > 0$: carat (weight)
- $p \in P = \{\text{I2}, \dots, \text{IF}\}$: clarity (lack of defects in crystal structure)
- $c \in C = \{J, \dots, D\}$: colour (how white the diamond is)
- $q \in Q = \{\text{Fair}, \dots, \text{Ideal}\}$: cut quality

Assume price is determined as follows:

$$v = x_0 + x_w\, w + \sum_{\hat{p} \in P \setminus \{p^\star\}} x_{\hat{p}}\, \delta_{p\hat{p}} + \sum_{\hat{c} \in C \setminus \{c^\star\}} x_{\hat{c}}\, \delta_{c\hat{c}} + \sum_{\hat{q} \in Q \setminus \{q^\star\}} x_{\hat{q}}\, \delta_{q\hat{q}}.$$

Parameters x can be determined through linear least squares fit.

See `9_least_squares_diamonds.jl`.

Least Squares

References and further reading

- L. N. Trefethen and D. Bau. *Numerical Linear Algebra.* Society for Industrial and Applied Mathematics (1997),
- J. W. Demmel. *Applied Numerical Linear Algebra.* Society for Industrial and Applied Mathematics (1997), `doi:10.1137/1.9781611971446`
- N. J. Higham. *Accuracy and Stability of Numerical Algorithms.* Society for Industrial and Applied Mathematics (2002), `doi:10.1137/1.9780898718027`