Least squares problems and QR factorization

Alternative way to solve linear least-squares problems: start from

$$A = QR, \quad Q = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix}, \quad R = \begin{bmatrix} R_1 \\ 0 \end{bmatrix}.$$

Trick: since orthogonal matrices preserve norm-2,

$$\begin{aligned} \|Ax - b\| &= \|Q^T(Ax - b)\| = \|Q^T QRx - Q^T b\| \\ &= \|Rx - Q^T b\| = \left\| \begin{bmatrix} R_1 \\ 0 \end{bmatrix} x - \begin{bmatrix} Q_1^T \\ Q_2^T \end{bmatrix} b \right\| \\ &= \left\| \begin{bmatrix} R_1 x - Q_1^T b \\ Q_2^T b \end{bmatrix} \right\|. \end{aligned}$$

Solving least squares with QR

$$\|Ax - b\| = \left\| \begin{bmatrix} R_1 x - Q_1^T b \\ Q_2^T b \end{bmatrix} \right\|$$

How can we minimize the norm of this vector?

Bottom block: same value, regardless of x. The squares of those entries will always be in the sum.

Top block I can choose x to make its entries smaller. Can I get $R_1 x - Q_1^T b = 0$? Yes, if R_1 invertible.

When is R_1 invertible?

Related to a result we have seen earlier. If $A = QR$, with Q orthogonal, then

$$A^T A = (QR)^T (QR) = R^T \underbrace{Q^T Q}_{=I} R = R^T R = \begin{bmatrix} R_1^T & 0 \end{bmatrix} \begin{bmatrix} R_1 \\ 0 \end{bmatrix} = R_1^T R_1.$$

A has full column rank $\iff$ $A^T A$ is posdef $\iff$ $A^T A = R_1^T R_1$ is invertible $\iff$ R_1 is inverbile.

(Note for the future: $R_1^T R_1$ is the Cholesky factorization of $A^T A$.)

Recap

If $A = QR = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \begin{bmatrix} R_1 \\ 0 \end{bmatrix}$ (and has full column rank), then the solution of $\min\|Ax - b\|$ is given by $x = R_1^{-1}(Q_1^T)b$.

Note that 'thin QR' $A = Q_1 R_1$ contains all we need here.

This gives (of course) the same solution x as the normal equations method. Indeed, some algebra gives

$$(A^T A)x = A^T b$$

$$A^T A = R_1^T R_1, \quad A^T b = R_1 Q_1^T b.$$

so the two methods solve similar linear systems:

Normal equations	QR-based method
$R_1^T R_1 x = R_1^T Q_1^T b$	$R_1 x = Q_1^T b.$

Exercises

1. Is it true that every symmetric, positive definite matrix $M \in \mathbb{R}^{m\times m}$ can be written as $B^T B$ for some $B \in \mathbb{R}^{m\times m}$? Hint: start from eigendecomposition $M = Q\Lambda Q^T$, and consider $\begin{bmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_m} \end{bmatrix}$.