Least squares by QR-factorization

Consider

$$y = Ax \qquad A_{m\times n} \text{ matrix with } m > n.$$

this system overdetermined. So we can't expect to solve the system exactly. Instead we solve it in least square sense

$$\min_x \|Ax - y\|$$

This system reduces to the normal equations:

$$A^T A x = A^T y.$$

if A is full column rank, then $A^T A$ is nonsingular and

$$x = (A^T A)^{-1} A^T y.$$

However using a matrix inverse to solve the system is more work and less accurate than solving the system with Gaussian elimination. But more importantly the normal equation are always more badly conditioned than the original overdetermined system. Infact the condition number is squared:

$$k(A^T A) = [k(A)]^2.$$

With finite-precision computation the normal equations can actually become singular and $(A^T A)^{-1}$ nonexistent even though the columns of A are independent.

As an extreme example consider:

$$A = \begin{pmatrix} 1 & 1 \\ \delta & 0 \\ 0 & \delta \end{pmatrix}$$

if δ is small, but nonzero the two columns of A are nearly parallel but are still linearly independent. The normal equations make the situation worse;

$$A^T A = \begin{pmatrix} 1+\delta^2 & 1 \\ 1 & 1+\delta^2 \end{pmatrix}.$$

if $|\delta| < 10^{-8}$, the matrix $A^T A$ computed with double-precision floating point arithmetic is exactly singular.

MATLAB avoids the normal equations and most computation is done by QR factorization.

Comparing with LU-factorization. The LU-factorization not always exist, e.g.

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

is nonsingular yet has no LU factorization unless rows are interchanged, whereas

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

is singular yet has LU factorization.

Another problem with LU factorization is we may loss information about the matrix A, e.g.

$$A = \begin{bmatrix} \varepsilon & 1 \\ 1 & 1 \end{bmatrix}$$ where ε is positive number.

Using LU factorization if rows are not interchanged we get

$$L = \begin{bmatrix} 1 & 0 \\ 1/\varepsilon & 1 \end{bmatrix} \text{ and } U = \begin{bmatrix} \varepsilon & 1 \\ 0 & -1/\varepsilon \end{bmatrix}.$$

But $LU = \begin{bmatrix} \varepsilon & 1 \\ 1 & 0 \end{bmatrix} \neq A$.

We summarize

(1) The LU-fact is the complete road map of Gaussian elimination applied to square nonsingular matrix whereas QR is the complete road map of Gram-Schmidt applied to a matrix with linearly independent columns.

(2) When they exist, both factorizations $A=LU$ and $A=QR$ are uniquely determined. by A.

(3) Once the LU factors exist of a nonsingular matrix A, the solution of $Ax=b$ is easily computed by- solve $Ly=b$ by forward substitution and then solve $Ux=y$ by back substitution. The QR factors can be used in a similar manner. If $A \in \mathbb{R}^{n\times n}$ is nonsingular, then $Q^T = Q^{-1}$, so $Ax=b \longleftrightarrow QRx=b \longleftrightarrow Rx=Q^Tb$, which is also a triangular system that is solved by back substitution.

(4) Unfortunately the LU factors of A doesn't exist when A is rectangular. But QR factors of A always exist as long as A has linearly independent columns.