Recall an $m \times n$ matrix A with linearly independent columns, can be written $A = QR$, where

- Q ← an $m \times n$ matrix has orthonormal columns and
- R ← an $n \times n$ matrix. is upper triangular.

You can compute this decomposition by

- finding an orthonormal basis for Col A, by applying Gram-Schmidt, and normalising, then using these vectors as the columns of Q.
- calculating $R = (IR = Q^T Q R =) Q^T A$.

Why? Solving $Ax=b$ is easier once you have a QR decomposition:

$$\begin{aligned} &Ax=b \\ &\Updownarrow \\ &QRx=b \\ &\Downarrow \\ &Q^TQRx=Q^Tb \\ &\Updownarrow \\ &Rx=Q^Tb \end{aligned}$$

which can be solved just by back substition — R is upper triangular!

Even better, this is less prone to numerical (rounding) errors than Gaussian elimination.

Today: Least squares solutions.

Sometimes $Ax=b$ doesn't have a solution, but we wish it did!

Perhaps the 'real' equations are $A'x=b'$, but we've made a small error, writing A instead of A' and b instead of b'!

A **least squares solution** to $Ax=b$ is an $\hat{x}$ so $A\hat{x}$ is as close as possible to b.

$\Longrightarrow$ This goes back to Gauss, who invented the method while deducing the orbit of the asteroid Ceres from just 3 observations!

Definition a least squares solution to $Ax=b$ is a vector $\hat{x}$ so

$$\|A\hat{x}-b\| \leq \|Ax-b\| \text{ for all vectors } x.$$

Recall Ax is always in the column space of A.

The equation $Ax=b$ is consistent if and only if $b \in \mathrm{Col}A$.

To find a least squares solution, we want Ax to be the closest point in $\mathrm{Col}A$ to b.

Thus the solution to the least squares problem is given by $\boxed{A\hat{x} = \mathrm{proj}_{\mathrm{Col}A}\, b.}$

Consider the QR decomposition $A = QR$.

Then $\text{proj}_{\text{Col} A} = \text{proj}_{\text{Col} Q} = QQ^T$, so we have

$$A\hat{x} = QR\hat{x} = QQ^T b$$

Multiplying by A^T, we have
$$\begin{aligned} A^T A\hat{x} &= R^T Q^T Q Q^T b \\ &= R^T Q^T b \\ &= A^T b. \end{aligned}$$

These are called the normal equations for a least squares solution:

$$\boxed{A^T A\hat{x} = A^T b}$$

(and don't require you to know anything by QR decompositions)

Examples

Find the least squares solution to

$$\begin{bmatrix} 1 & 3 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} x = \begin{bmatrix} 5 \\ 1 \\ 0 \end{bmatrix}.$$

Let $A = \begin{bmatrix} 1 & 5 \\ 1 & -1 \\ 1 & 1 \end{bmatrix}$. We'll solve the normal equations,

$$A^T A \hat{x} = A^T b.$$

$$A^T A = \begin{bmatrix} 1 & 1 & 1 \\ 3 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ 3 & 11 \end{bmatrix}$$

$A^+ b = \begin{bmatrix} 6 \\ 14 \end{bmatrix}$. We need to solve $\begin{bmatrix} 3 & 3 \\ 3 & 11 \end{bmatrix} \hat{x} = \begin{bmatrix} 6 \\ 14 \end{bmatrix}$.

Some row-reduction on the augmented matrix:

$$\left[\begin{array}{cc|c} 3 & 3 & 6 \\ 3 & 11 & 14 \end{array}\right] \to \left[\begin{array}{cc|c} 3 & 3 & 6 \\ 0 & 8 & 8 \end{array}\right] \to \left[\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & 1 & 1 \end{array}\right]$$

$$\to \left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 1 \end{array}\right], \text{ so } \hat{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Note that $A\hat{x} = \begin{bmatrix} 1 & 3 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ 2 \end{bmatrix}$ is the closest point in Col A to $\begin{bmatrix} 5 \\ 1 \\ 0 \end{bmatrix}$.

Example

Find the least squares solution to

$$Ax=b,$$

where

$$A=\begin{bmatrix}3 & -1\\ 1 & -2\\ 2 & 3\end{bmatrix},$$

$$b=\begin{bmatrix}4\\ 3\\ 2\end{bmatrix}.$$

$$A^TA=\begin{bmatrix}3 & 1 & 2\\ -1 & -2 & 3\end{bmatrix}\begin{bmatrix}3 & -1\\ 1 & -2\\ 2 & 3\end{bmatrix}=\begin{bmatrix}14 & 1\\ 1 & 14\end{bmatrix}.$$

This is invertible — we can rearrange the normal equations as

$$\hat{x}=(A^TA)^{-1}A^Tb$$

$$=\begin{bmatrix}14 & 1\\ 1 & 14\end{bmatrix}^{-1}\begin{bmatrix}3 & 1 & 2\\ -1 & -2 & 3\end{bmatrix}\begin{bmatrix}4\\ 3\\ 2\end{bmatrix}$$

$$=\frac{1}{195}\begin{bmatrix}14 & -1\\ -1 & 14\end{bmatrix}\begin{bmatrix}19\\ -4\end{bmatrix}$$

$$=\frac{1}{13}\begin{bmatrix}18\\ -5\end{bmatrix}.$$

Example

Find the least squares solutions to $Ax=b$, with

$$A=\begin{bmatrix}1&0&2\\2&1&5\\-1&1&-1\\0&1&1\end{bmatrix}$$

$$b=\begin{bmatrix}1\\-1\\-1\\2\end{bmatrix}.$$

$$A^TA=\begin{bmatrix}6&1&13\\1&3&5\\13&5&31\end{bmatrix},\quad A^Tb=\begin{bmatrix}0\\0\\0\end{bmatrix}.$$

Thus solving $A^TA\hat{x}=A^Tb$ is the same as finding $\mathrm{Nul}(A^TA)$.

$$\begin{bmatrix}6&1&13\\1&3&5\\13&5&31\end{bmatrix}\xrightarrow{\text{rref}}\begin{bmatrix}1&0&2\\0&1&1\\0&0&0\end{bmatrix},$$

So $\hat{x}\in \mathrm{Nul}\,A^TA=\mathrm{span}\left\{\begin{bmatrix}2\\1\\-1\end{bmatrix}\right\}.$

The solution isn't unique.

There is a unique least squares solution if and only if A^TA is invertible, or equivalently, the columns of A are independent.

Actually, if we have a QR decomposition, we can do much better:

$$A\hat{x} = QR\hat{x} = QQ^{T}b$$

$$\Downarrow \text{ (multiplying by } Q^{T})$$

$$\begin{array}{ccc} Q^{T}QR\hat{x} \cancel{= R\hat{x}} & = & Q^{T}QQ^{T}b \\ \| & & \\ R\hat{x} & = & Q^{T}b \end{array}$$

in practice, just solve this by back-substitution, rather than calculating R^{-1}.

so $\hat{x} = R^{-1}Q^{T}b$

Even better, this is much less susceptible to rounding errors than solving the normal equations.

Example

Suppose

$$A = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 0 & 5 \end{bmatrix}$$

$$b = \begin{bmatrix} -1 \\ 6 \\ 5 \\ 7 \end{bmatrix}.$$

Find the least squares solution to

$$Ax = b$$

Since we've been given a QR decomposition of A, let's use the equations

$$R\hat{x} = Q^T b.$$

$$Q^T b = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} -1 \\ 6 \\ 5 \\ 7 \end{bmatrix} = \begin{bmatrix} \frac{17}{2} \\ \frac{9}{2} \end{bmatrix}.$$

Then solve

$$\begin{bmatrix} 2 & 3 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{17}{2} \\ \frac{9}{2} \end{bmatrix}$$

So $x_2 = \frac{9}{10}$, $2x_1 + 3x_2 = \frac{17}{2}$

$$x_1 = \frac{17}{4} - \frac{27}{20} = \frac{85-27}{20} = \frac{58}{20} = \frac{29}{10}.$$

A cautionary example

Different methods matter, when doing numerical linear algebra.

$$A = \begin{pmatrix} .780 & .563 \\ .913 & .659 \end{pmatrix} \qquad b = \begin{pmatrix} .217 \\ .254 \end{pmatrix}$$

(from http://www.cs.toronto.edu/~enright/teaching/D37/Week2.pdf)

$$A^T A = \begin{pmatrix} 1.44 & 1.04 \\ 1.04 & 0.751 \end{pmatrix}$$

$$A^T b = \begin{pmatrix} .312 \\ .366 \end{pmatrix}$$

$$\left(\begin{array}{cc|c} 1.44 & 1.04 & .312 \\ 1.04 & 0.751 & .366 \end{array}\right) \longrightarrow \left(\begin{array}{cc|c} 1.44 & 1.04 & .312 \\ 0 & -0.000111 & .141 \end{array}\right)$$

$$\Rightarrow x_2 = -1270$$

$$x_1 = 917$$

$$Ax = \begin{pmatrix} -474 \\ -555 \end{pmatrix}$$

, (1,-1) is an exact solution!

ul using this example in future years; it was done in a hurry...)

Alternatively:

$$Q = \begin{pmatrix} -0.650 & -0.760 \\ -0.760 & 0.650 \end{pmatrix}$$

$$R = \begin{pmatrix} -1.20 & -0.867 \\ 0 & 8.33 \times 10^{-7} \end{pmatrix}$$

$$\hat{x} = R^{-1} Q^T b = \begin{bmatrix} -156 \\ 216 \end{bmatrix}$$

$$A\hat{x} = \begin{bmatrix} 0.239 \\ 0.280 \end{bmatrix}$$