The Gram–Schmidt Process

Given a basis $\{\mathbf{x}_1, \ldots, \mathbf{x}_p\}$ for a nonzero subspace W of $\mathbb{R}^n$, define

$$\begin{aligned}
\mathbf{v}_1 &= \mathbf{x}_1 \\
\mathbf{v}_2 &= \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 \\
\mathbf{v}_3 &= \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 \\
&\vdots \\
\mathbf{v}_p &= \mathbf{x}_p - \frac{\mathbf{x}_p \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_p \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 - \cdots - \frac{\mathbf{x}_p \cdot \mathbf{v}_{p-1}}{\mathbf{v}_{p-1} \cdot \mathbf{v}_{p-1}}\mathbf{v}_{p-1}
\end{aligned}$$

Then $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ is an orthogonal basis for W. In addition

$$\operatorname{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} = \operatorname{Span}\{\mathbf{x}_1, \ldots, \mathbf{x}_k\} \qquad \text{for } 1 \le k \le p$$

EXAMPLE 1 Let $W = \text{Span}\{\mathbf{x}_1, \mathbf{x}_2\}$, where $\mathbf{x}_1 = \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix}$ and $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$. Construct an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2\}$ for W.

EXAMPLE 2 Let $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, and $\mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$. Then $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ is clearly linearly independent and thus is a basis for a subspace W of $\mathbb{R}^4$. Construct an orthogonal basis for W.

The QR Factorization

If A is an $m \times n$ matrix with linearly independent columns, then A can be factored as $A = QR$, where Q is an $m \times n$ matrix whose columns form an orthonormal basis for Col A and R is an $n \times n$ upper triangular invertible matrix with positive entries on its diagonal.

Example:

Find the QR decomposition of the following matrix

$$A = \begin{bmatrix} 5 & 9 \\ 1 & 7 \\ -3 & -5 \\ 1 & 5 \end{bmatrix}$$

Solution:

Since A has independent column, then by QR decomposition, we can decompose A into product of two matrices,

$$A = QR,$$

where, Q is orthonormal matrix (orthonormal basis for the column space of A),

R is upper triangular matrix with positive entries of the diagonal.

Let us calculate the matrix Q.

Before, let us calculate orthogonal basis of the column space of A.

By using Gram-Schmidt orthogonalization process:

$\{u_1, u_2\}$ is the orthogonal basis of the column space of A.

$$u_1 = x_1, u_2 = x_2 - \left(\frac{x_2 \cdot u_1}{u_1 \cdot u_1}\right) u_1$$

$$x_1 = \begin{bmatrix} 5 \\ 1 \\ -3 \\ 1 \end{bmatrix}, x_2 = \begin{bmatrix} 9 \\ 7 \\ -5 \\ 5 \end{bmatrix}$$

$$u_1 = x_1 = \begin{bmatrix} 5 \\ 1 \\ -3 \\ 1 \end{bmatrix}$$

$$x_2 . u_1 = \begin{bmatrix} 9 \\ 7 \\ -5 \\ 5 \end{bmatrix} \cdot \begin{bmatrix} 5 \\ 1 \\ -3 \\ 1 \end{bmatrix} = 45 + 7 + 15 + 5 = 72$$

$$u_1 \cdot u_1 = \begin{bmatrix} 5 \\ 1 \\ -3 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 5 \\ 1 \\ -3 \\ 1 \end{bmatrix} = 25 + 1 + 9 + 1 = 36$$

$$u_2 = \begin{bmatrix} 9 \\ 7 \\ -5 \\ 5 \end{bmatrix} - \left(\frac{72}{36}\right) \begin{bmatrix} 5 \\ 1 \\ -3 \\ 1 \end{bmatrix} = \begin{bmatrix} 9 \\ 7 \\ -5 \\ 5 \end{bmatrix} - 2 \begin{bmatrix} 5 \\ 1 \\ -3 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 5 \\ 1 \\ 3 \end{bmatrix}$$

Orthogonal basis$= \left\{ \begin{bmatrix} 5 \\ 1 \\ -3 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 5 \\ 1 \\ 3 \end{bmatrix} \right\}$

Orthonormal vectors are $v_1 = \frac{u_1}{\|u_1\|}, v_2 = \frac{u_2}{\|u_2\|}$

$\|u_1\| = \sqrt{25 + 1 + 9 + 1} = 6$

$\|u_2\| = \sqrt{1 + 25 + 1 + 9} = 6$

$$v_1 = \frac{1}{6}\begin{bmatrix} 5 \\ 1 \\ -3 \\ 1 \end{bmatrix}, v_2 = \frac{1}{6}\begin{bmatrix} -1 \\ 5 \\ 1 \\ 3 \end{bmatrix}$$

$\{v_1, v_2\}$ forms an orthonormal basis.

Therefore, $Q = \begin{bmatrix} \frac{5}{6} & -\frac{1}{6} \\ \frac{1}{6} & \frac{5}{6} \\ -\frac{3}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{3}{6} \end{bmatrix}$.

Since, we have $A = QR$,

Multiplying Q^T on both sides, we have

$Q^T A = Q^T QR \Rightarrow R = Q^T A$

$$R = Q^T A$$

$$R = \begin{bmatrix} \frac{5}{6} & \frac{1}{6} & -\frac{3}{6} & \frac{1}{6} \\ -\frac{1}{6} & \frac{5}{6} & \frac{1}{6} & \frac{3}{6} \end{bmatrix} \begin{bmatrix} 5 & 9 \\ 1 & 7 \\ -3 & -5 \\ 1 & 5 \end{bmatrix} = \begin{bmatrix} 6 & 12 \\ 0 & 6 \end{bmatrix}$$

Therefore, $\begin{bmatrix} 5 & 9 \\ 1 & 7 \\ -3 & -5 \\ 1 & 5 \end{bmatrix} = \begin{bmatrix} \frac{5}{6} & -\frac{1}{6} \\ \frac{1}{6} & \frac{5}{6} \\ -\frac{3}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{3}{6} \end{bmatrix} \begin{bmatrix} 6 & 12 \\ 0 & 6 \end{bmatrix}$.

Example:

Find a QR factorization of $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$.

Solution:
Home work.

Hint:
$\{u_1, u_2, u_3\}$ is the orthogonal basis for the column space of the matrix A.

$A = [x_1 \; x_2 \; x_3]$

$$x_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, x_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, x_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$$

$$u_1 = x_1$$

$$u_2 = x_2 - \left(\frac{x_2 \cdot u_1}{u_1 \cdot u_1}\right) u_1$$

$$u_3 = x_3 - \left(\frac{x_3 \cdot u_1}{u_1 \cdot u_1}\right) u_1 - \left(\frac{x_3 \cdot u_2}{u_2 \cdot u_2}\right) u_2$$

do the rest ....

Note:

1) To have solution for the system $A\mathbf{x} = \mathbf{b}$, $\boldsymbol{b}$ must be in the column space of A.

Proof: For example consider $A = \begin{bmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ a_7 & a_8 & a_9 \end{bmatrix}$, $b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$, and $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ is the solution of the system.$A\mathbf{x} = \mathbf{b}$

$$\Rightarrow \begin{bmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ a_7 & a_8 & a_9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

It can be written as follows

$$\begin{bmatrix} a_1 \\ a_4 \\ a_7 \end{bmatrix} x_1 + \begin{bmatrix} a_2 \\ a_5 \\ a_8 \end{bmatrix} x_2 + \begin{bmatrix} a_3 \\ a_6 \\ a_9 \end{bmatrix} x_3 = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

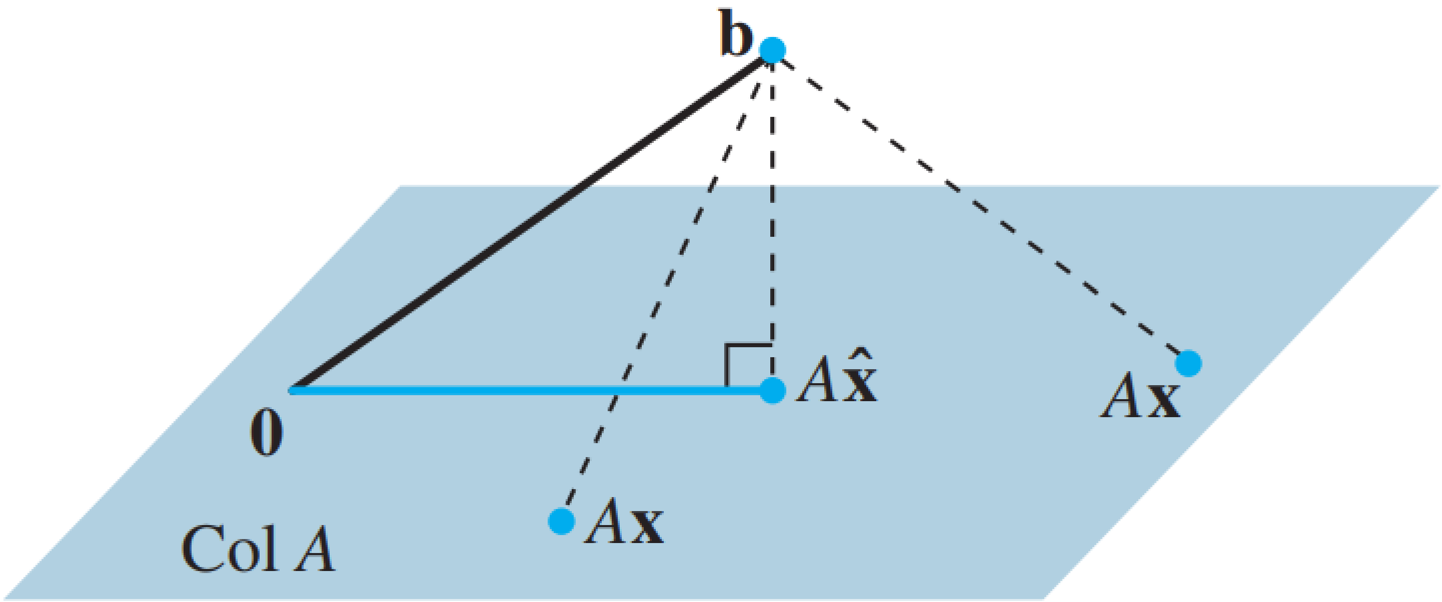
$\Rightarrow c_1x_1 + c_2x_2 + c_3x_3 = b$, where c_1, c_2, c_3 are the columns of the matrix A.

From the expression it is clear that the matrix $\boldsymbol{b}$ expressed as the linear combination of the columns of the matrix A. That is, we can say that $\boldsymbol{b}$ is in the column space of A.

Least Squares solution:

Definition:

If A is $m \times n$ and b is in R^m, a least-squares solution of $Ax = b$ is an $\hat{x}$ in R^n such that $||b - A\hat{x}|| \leq ||b - Ax||$ for all x in R^n.



Least Squares: Four Ways:

1 The SVD of A leads to its **pseudoinverse** $\boldsymbol{A}^{+}$. Then $\widehat{\boldsymbol{x}} = \boldsymbol{A}^{+}\boldsymbol{b}$: One short formula.

2 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\widehat{\boldsymbol{x}} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$ can be solved directly when A has **independent columns**.

3 The Gram-Schmidt idea produces **orthogonal columns** in Q. Then $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{R}$.

4 **Minimize** $||\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}||^2 + \boldsymbol{\delta}^2||\boldsymbol{x}||^2$. That penalty changes the normal equations to $(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} + \boldsymbol{\delta}^2\boldsymbol{I})\boldsymbol{x}_{\delta} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$. Now the matrix is invertible and $\boldsymbol{x}_{\delta}$ goes to $\widehat{\boldsymbol{x}}$ as $\delta \to 0$.

The Least Squares solution to $Ax = b$ is $x^+ = A^+b$

1) $x = x^+ = A^+b$ makes $||b - Ax||^2$ as small as possible. Least Squares solution
2) If another $\hat{x}$ achieves that minimum then $||x^+|| < ||\hat{x}||$. Minimum norm solution

x^+ is the minimum norm least square solution.

Normal equation for $\widehat{x}$	$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\,\widehat{\boldsymbol{x}} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$
Least squares solution to $Ax = \boldsymbol{b}$	$\widehat{\boldsymbol{x}} = (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$
Projection of $\boldsymbol{b}$ onto the column space of A	$\boldsymbol{p} = \boldsymbol{A}\,\widehat{\boldsymbol{x}} = \boldsymbol{A}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$
Projection matrix that multiplies $\boldsymbol{b}$ to give $\boldsymbol{p}$	$\boldsymbol{P} = \boldsymbol{A}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}$

Example:

Find a least-squares solution of $A\mathbf{x} = \mathbf{b}$ for

$$A = \begin{bmatrix} 1 & -6 \\ 1 & -2 \\ 1 & 1 \\ 1 & 7 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} -1 \\ 2 \\ 1 \\ 6 \end{bmatrix}$$

Also, determine the least-squared error in the least square solution of the given system.

Solution:

Clearly is it can be verified that $\boldsymbol{b}$ is not in the column space of A.
Therefore, $A\mathbf{x} = \mathbf{b}$ is inconsistent.

We can find the least square solution for the given system.

First we will take the projection of $\boldsymbol{b}$ on to the column space of A. We denote that with $\widehat{\boldsymbol{b}}$.

Let c_1, c_2 are the columns of the matrix A.

$\hat{b}$ can be written as $\hat{b} = Proj_A b$.

$$\hat{b} = \left(\frac{b \cdot c_1}{c_1 \cdot c_1}\right) c_1 + \left(\frac{b \cdot c_2}{c_2 \cdot c_2}\right) c_2$$

$$b = \begin{bmatrix} -1 \\ 2 \\ 1 \\ 6 \end{bmatrix}, c_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, c_2 = \begin{bmatrix} -6 \\ -2 \\ 1 \\ 7 \end{bmatrix}$$

$$b \cdot c_1 = \begin{bmatrix} -1 \\ 2 \\ 1 \\ 6 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = -1 + 2 + 1 + 6 = 8; \qquad c_1 \cdot c_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = 1 + 1 + 1 + 1 = 4$$

$$b \cdot c_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \\ 6 \end{bmatrix} \cdot \begin{bmatrix} -6 \\ -2 \\ 1 \\ 7 \end{bmatrix} = 6 - 4 + 1 + 42 = 45; \qquad c_2 \cdot c_2 = \begin{bmatrix} -6 \\ -2 \\ 1 \\ 7 \end{bmatrix} \cdot \begin{bmatrix} -6 \\ -2 \\ 1 \\ 7 \end{bmatrix} = 36 + 4 + 1 + 49 = 90$$

$$\hat{b} = \frac{8}{4}\begin{bmatrix}1\\1\\1\\1\end{bmatrix} + \frac{90}{45}\begin{bmatrix}-6\\-2\\1\\7\end{bmatrix} = 2\begin{bmatrix}1\\1\\1\\1\end{bmatrix} + 2\begin{bmatrix}-6\\-2\\1\\7\end{bmatrix} = \begin{bmatrix}-10\\-2\\4\\16\end{bmatrix}$$

Now, solve $A\mathbf{x} = \widehat{\boldsymbol{b}}$

$$\begin{bmatrix}1 & -6\\1 & -2\\1 & 1\\1 & 7\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} = \begin{bmatrix}-10\\-2\\4\\16\end{bmatrix}$$

Augmented matrix $\begin{bmatrix}1 & -6 & -10\\1 & -2 & -2\\1 & 1 & 4\\1 & 7 & 16\end{bmatrix}$

By Gauss elimination method we need to make the elements below the principal diagonals to zeros.

$R_2 \to R_2 - R_1; R_3 \to R_3 - R_1; R_4 \to R_4 - R_1$

$$\begin{bmatrix} 1 & -6 & -10 \\ 0 & 4 & 8 \\ 0 & 7 & 14 \\ 0 & 13 & 26 \end{bmatrix} \Rightarrow \textit{It is clear that } R_2, R_3, R_4 \text{ are dependent.}$$

Therefore, we have only two L.I. equations.

Which are, $x_1 - 6x_2 = -10$,

$$4x_2 = 8 \Rightarrow x_2 = 2$$

$$7x_2 = 14 \Rightarrow x_2 = 2$$

$$13x_2 = 26 \Rightarrow x_2 = 2$$

$$x_1 = -10 + 6(2) = 2$$

Therefore, the least square solution is $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$

The Least square error is $\|b - A\hat{x}\| = \left\| \begin{bmatrix} -1 \\ 2 \\ 1 \\ 6 \end{bmatrix} - \begin{bmatrix} -10 \\ -2 \\ 4 \\ 16 \end{bmatrix} \right\| = \left\| \begin{bmatrix} 9 \\ 4 \\ -3 \\ -10 \end{bmatrix} \right\| = \sqrt{49 + 16 + 9 + 100} = \sqrt{174}$.

Example:

Find the least-squares solution of $A\mathbf{x} = \mathbf{b}$ for

$$A = \begin{bmatrix} 1 & 3 & 5 \\ 1 & 1 & 0 \\ 1 & 1 & 2 \\ 1 & 3 & 3 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 3 \\ 5 \\ 7 \\ -3 \end{bmatrix}$$ using QR decomposition

Example:

Find a least-squares solution of the inconsistent system $A\mathbf{x} = \mathbf{b}$ for

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$$ using QR factorization.

Also, determine the least-squared error in the least square solution of the given system.

Solution:

Given $A\mathbf{x} = \mathbf{b}$ is a inconsistent system.

Hint:

Since A has independent columns. Therefore, we can decompose A into

$$A = QR$$

Where, Q is orthonormal matrix (columns of Q forms an orthonormal basis of the column space of A) and R is the upper triangular matrix with positive diagonal elements.

We can find the columns of Q by Grams-Schmidt orthogonalization process.

After finding Q, we can find R by $R = Q^T A$

Given $A\mathbf{x} = \mathbf{b} \Rightarrow (\mathrm{QR})\mathbf{x} = \mathbf{b} \Rightarrow \mathrm{Rx} = \mathrm{Q}^{\mathrm{T}}\mathbf{b}$.

By back substitution, we can find the x.

Normal Equations:

Let $A\mathbf{x} = \mathbf{b}$ be the given system then $A^T A\mathrm{x} = \mathrm{A}^{\mathrm{T}}\mathbf{b}$ is called the normal equations.

Example:

Find a least-squares solution of the inconsistent system $A\mathbf{x} = \mathbf{b}$ for

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$$

Solution:

From the normal equations $A^T A\mathrm{x} = \mathrm{A}^{\mathrm{T}}\boldsymbol{b}$

$$A^T A = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix}$$

$$A^T \mathbf{b} = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$

Then the equation $A^T A\mathbf{x} = A^T\mathbf{b}$ becomes

$$\begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$

Then the equation $A^TA\mathbf{x} = A^T\mathbf{b}$ becomes

$$\begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$

Row operations can be used to solve this system, but since A^TA is invertible and 2×2, it is probably faster to compute

$$(A^TA)^{-1} = \frac{1}{84}\begin{bmatrix} 5 & -1 \\ -1 & 17 \end{bmatrix}$$

and then to solve $A^TA\mathbf{x} = A^T\mathbf{b}$ as

$$\begin{aligned} \hat{\mathbf{x}} &= (A^TA)^{-1}A^T\mathbf{b} \\ &= \frac{1}{84}\begin{bmatrix} 5 & -1 \\ -1 & 17 \end{bmatrix}\begin{bmatrix} 19 \\ 11 \end{bmatrix} = \frac{1}{84}\begin{bmatrix} 84 \\ 168 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \end{aligned} \qquad \blacksquare$$

Example:

Find a least-squares solution of $A\mathbf{x} = \mathbf{b}$ for

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} -3 \\ -1 \\ 0 \\ 2 \\ 5 \\ 1 \end{bmatrix}$$ using normal equations.

Also, determine the least-squared error in the least square solution of the given system.

Theorem:

Let A be an $m \times n$ matrix. The following statements are logically equivalent:

(a) The equation $Ax = b$ has a unique least-squares solution for each b in $\mathbb{R}^m$.
(b) The columns of A are linearly independent.
(c) The matrix $A^T A$ is invertible.

When these statements are true, the least-squares solution $\hat{\mathrm{x}}$ is given by

$$\hat{\mathrm{x}} = (A^T A)^{-1} A^T \boldsymbol{b}$$