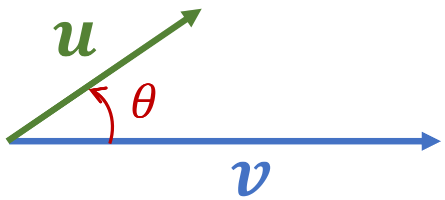
ECSE 343 Numerical Methods in Engineering

Roni Khazaka
Dept. of Electrical and Computer Engineering
McGill University

Inner Product

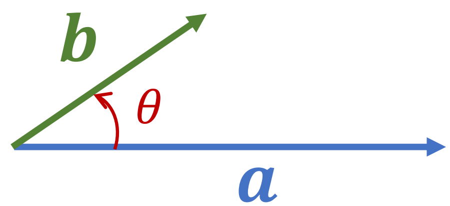


$$\langle u, v \rangle = |u||v|\cos(\theta)$$

$$u = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \qquad v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

$$\langle u, v \rangle = \sum_{i=1}^{n} u_i v_i$$

Inner Product

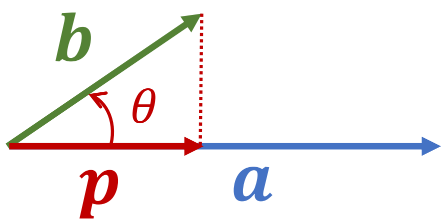


$$\langle a, b \rangle = |a||b|\cos(\theta)$$

$$b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \qquad a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

$$\langle a, b \rangle = \sum_{i=1}^{n} a_i b_i = a^T b$$

Orthogonal Projection



$$|p| = |b|\cos(\theta)$$

$$\langle a, b\rangle = |a||b|\cos(\theta)$$

$$|p| = \frac{\langle a, b\rangle}{|a|} = \frac{a^T b}{|a|}$$

$$p = \hat{a}|p|$$

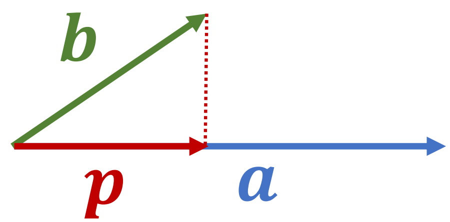
$\hat{a}$ is the unit vector in direction of a

$$\hat{a} = \frac{a}{|a|}$$

$$p = \frac{a}{|a|}\frac{a^T b}{|a|} = \frac{aa^T}{|a||a|}b$$

$$p = \frac{aa^T}{a^T a}b$$

Orthogonal Projection



$$\boldsymbol{p} = \frac{\boldsymbol{a}\boldsymbol{a}^T}{\boldsymbol{a}^T\boldsymbol{a}}\boldsymbol{b}$$

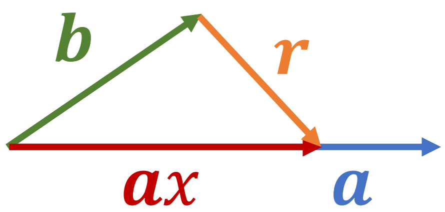
$$\boldsymbol{p} = \boldsymbol{P}\boldsymbol{b}$$

$$a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

$$\boldsymbol{P} = \frac{\boldsymbol{a}\boldsymbol{a}^T}{\boldsymbol{a}^T\boldsymbol{a}}$$

Projection Matrix

Solution of One Equation



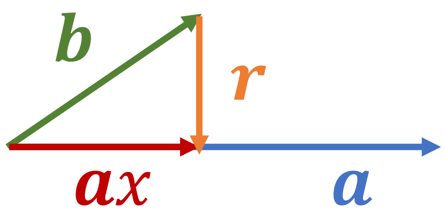
Find scalar x such that: $\boldsymbol{a}x = \boldsymbol{b}$

Impossible: $\boldsymbol{a}x \neq \boldsymbol{b}$ for all x

Minimize residual $\boldsymbol{r} = \boldsymbol{a}x - \boldsymbol{b}$

i.e. minimize $\|\boldsymbol{r}\| = \|\boldsymbol{a}x - \boldsymbol{b}\|$

Solution of One Equation



Find scalar x such that: $\boldsymbol{a}x = \boldsymbol{b}$

Impossible: $\boldsymbol{a}x \neq \boldsymbol{b}$ for all x

Minimize $\boldsymbol{r} = \boldsymbol{a}x - \boldsymbol{b}$

i.e. minimize $||\boldsymbol{r}|| = ||\boldsymbol{a}x - \boldsymbol{b}||$

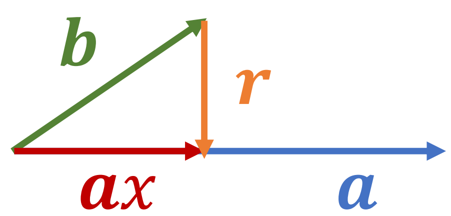
$||\boldsymbol{r}||$ is minimum when $\boldsymbol{r} \perp \boldsymbol{a}$

$||\boldsymbol{r}||$ is minimum when $\boldsymbol{a}^T\boldsymbol{r} = \boldsymbol{0}$

$||\boldsymbol{r}||$ is minimum when $\boldsymbol{a}^T(\boldsymbol{a}x - \boldsymbol{b}) = \boldsymbol{0}$

$||\boldsymbol{r}||$ is minimum when $\boxed{\boldsymbol{a}^T\boldsymbol{a}x = \boldsymbol{a}^T\boldsymbol{b}}$

Solution of One Equation



$\|r\|$ is minimum when $x = \dfrac{a^T b}{a^T a}$

$\|r\|$ is minimum when ax is the orthogonal projection of b on a

The orthogonal projection of b on a is

$$p = ax = a\frac{a^T b}{a^T a} = \frac{aa^T}{a^T a} b$$

Overdetermined System

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \\ a_{51} & a_{52} & a_{53} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{bmatrix}$$

- 3 degrees of freedom (unknowns)
- 5 constraints (equations)
- Solution may or may not exist.

Least squares method: Can find approximate solution by minimizing the residual: $||\boldsymbol{Ax} - \boldsymbol{b}||$

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

Overdetermined Systems

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \\ a_{51} & a_{52} & a_{53} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{bmatrix}$$

- Assume columns of A are linearly independent
- Columns of A form a 3D subspace Q of $\mathbb{R}^5$

Overdetermined Systems

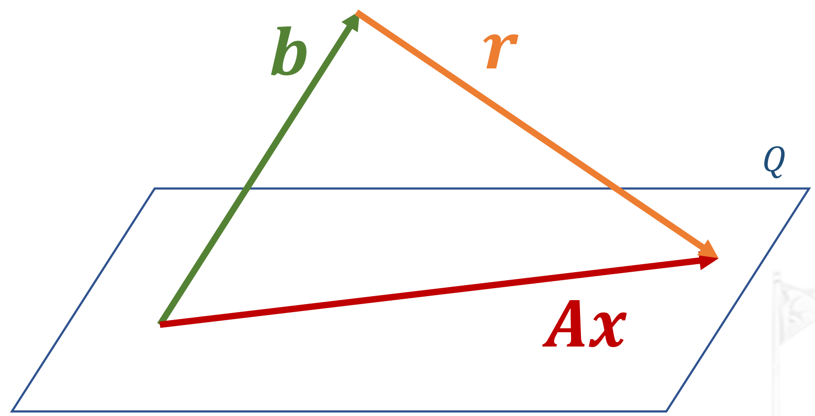
Find $\boldsymbol{x}$ such that: $\boldsymbol{Ax} = \boldsymbol{b}$

Impossible:

$\boldsymbol{Ax} \in Q \qquad \boldsymbol{b} \notin Q$

Minimize $\boldsymbol{r} = \boldsymbol{Ax} - \boldsymbol{b}$

i.e. minimize $||\boldsymbol{r}|| = ||\boldsymbol{Ax} - \boldsymbol{b}||$



Overdetermined Systems

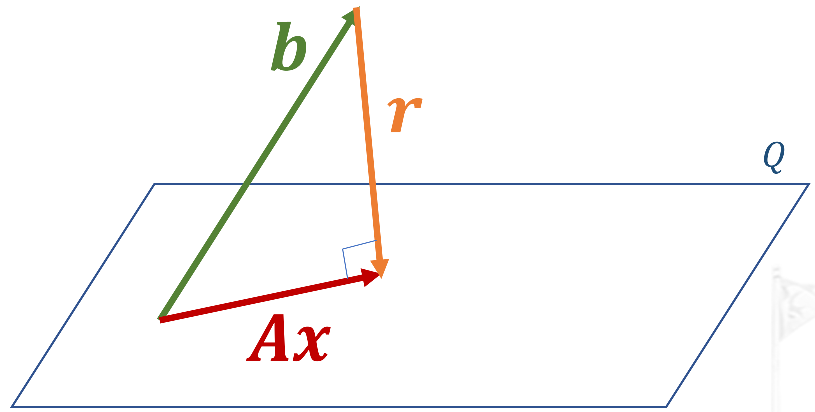
Minimize $||\boldsymbol{r}|| = ||\boldsymbol{Ax} - \boldsymbol{b}||$

$\Longrightarrow \boldsymbol{r} \perp Q$

$\Longrightarrow \boldsymbol{A}^T(\boldsymbol{Ax} - \boldsymbol{b}) = \boldsymbol{0}$

$\Longrightarrow \boxed{\boldsymbol{A}^T\boldsymbol{Ax} = \boldsymbol{A}^T\boldsymbol{b}}$

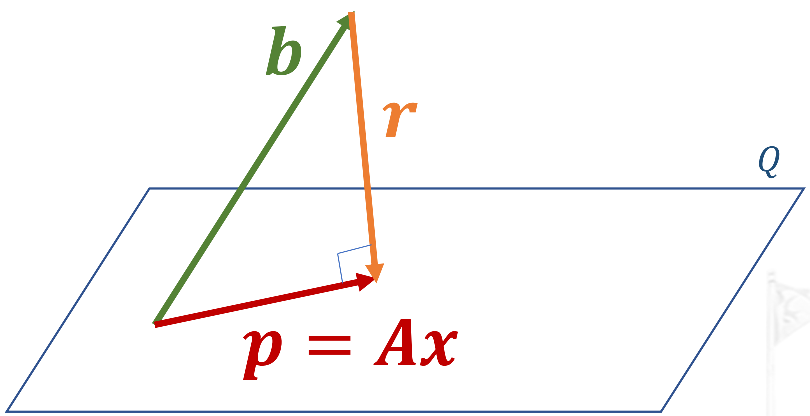
$\Longrightarrow \boldsymbol{x} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{b}$



Projection onto Q

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

$$\boldsymbol{p} = \boldsymbol{A}\boldsymbol{x} = \underbrace{\boldsymbol{A}(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T}_{\text{Projection Matrix}}\boldsymbol{b}$$



Overdetermined System

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \\ a_{51} & a_{52} & a_{53} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{bmatrix}$$

- 3 degrees of freedom (unknowns)
- 5 constraints (equations)
- Solution may or may not exist.

Least squares method: Can find approximate solution by minimizing the residual: $||\boldsymbol{Ax} - \boldsymbol{b}||$

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

Least Squares Approximation

Consider n data points (t_i, y_i)

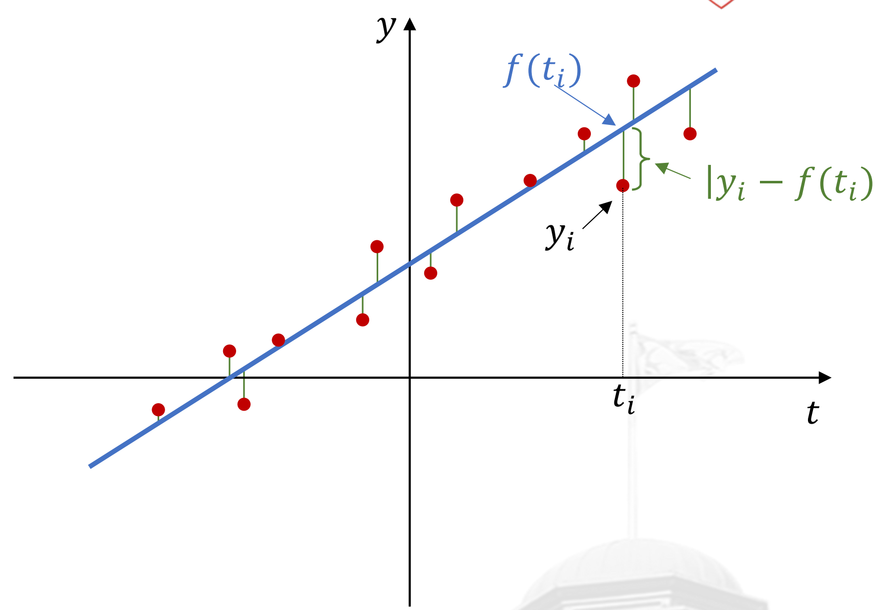
Approximate data with a model:

$$y = f(t) = a_o + a_1 t$$

a_o and a_1 are the model parameters.

Choose the parameters to minimize:

$$e = \sum_{i=1}^{n} (f(t_i) - y_i)^2$$



Least Squares Approximation

Consider n data points (t_i, y_i)

$$y = f(t) = a_o + a_1 t$$

Choose the parameters to minimize: $$e = \sum_{i=1}^{n} (f(t_i) - y_i)^2$$

Let: $$\boldsymbol{r} \equiv \begin{bmatrix} f(t_1) - y_1 \\ \vdots \\ f(t_i) - y_i \\ \vdots \\ f(t_n) - y_n \end{bmatrix}$$

Then: $$e = \boldsymbol{r}^T \boldsymbol{r} = \|\boldsymbol{r}\|_2^2$$

Least Squares Approximation

Consider n data points (t_i, y_i)

$$y = f(t) = a_o + a_1 t$$

$$\boldsymbol{r} = \begin{bmatrix} f(t_1) - y_1 \\ \vdots \\ f(t_i) - y_i \\ \vdots \\ f(t_n) - y_n \end{bmatrix} = \begin{bmatrix} a_o + a_1 t_1 - y_1 \\ \vdots \\ a_o + a_1 t_i - y_i \\ \vdots \\ a_o + a_1 t_n - y_n \end{bmatrix} = \begin{bmatrix} 1 & t_1 \\ \vdots & \vdots \\ 1 & t_i \\ \vdots & \vdots \\ 1 & t_n \end{bmatrix} \begin{bmatrix} a_o \\ a_1 \end{bmatrix} - \begin{bmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_n \end{bmatrix}$$

Minimize : $e = \boldsymbol{r}^T \boldsymbol{r} = \|\boldsymbol{r}\|_2^2$

Overdetermined System

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \\ a_{51} & a_{52} & a_{53} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{bmatrix}$$

- 3 degrees of freedom (unknowns)
- 5 constraints (equations)
- Solution may or may not exist.

Least squares method: Can find approximate solution by minimizing the residual: $||\boldsymbol{Ax} - \boldsymbol{b}||$

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

Least Squares Approximation

$$\begin{bmatrix} 1 & t_1 \\ \vdots & \vdots \\ 1 & t_i \\ \vdots & \vdots \\ 1 & t_n \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_n \end{bmatrix}$$

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

The parameters that result in a least squares approximation are the solution of:

$$\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{b}$$

Least Squares Approximation

Consider n data points (t_i, y_i)

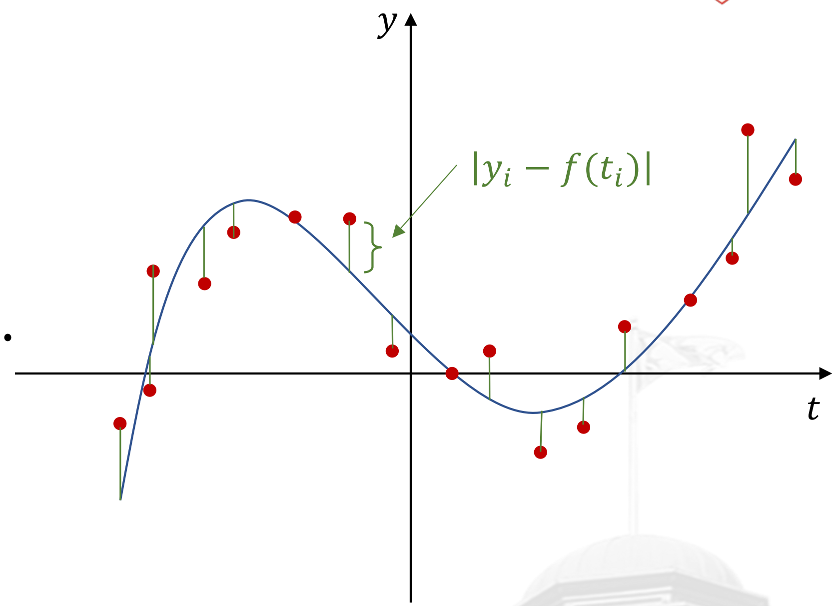
Approximate data with a model:

$$y = f(t) = a_o + a_1 t + a_2 t^2 + a_3 t^3 + \cdots$$

$a_o, a_1, \ldots$ are the model parameters.

Choose the parameters to minimize:

$$e = \sum_{i=1}^{n} (f(t_i) - y_i)^2$$



Least Squares Approximation

Consider n data points (t_i, y_i)

$$y = f(t) = a_o + a_1 t + a_2 t^2 + a_3 t^3 + \cdots$$

Choose the parameters to minimize: $$e = \sum_{i=1}^{n} (f(t_i) - y_i)^2$$

Let: $$\boldsymbol{r} \equiv \begin{bmatrix} f(t_1) - y_1 \\ \vdots \\ f(t_i) - y_i \\ \vdots \\ f(t_n) - y_n \end{bmatrix}$$

Then: $$e = \boldsymbol{r}^T \boldsymbol{r} = \|\boldsymbol{r}\|_2^2$$

Least Squares Approximation

Consider n data points (t_i, y_i)

$$y = f(t) = a_o + a_1 t + a_2 t^2 + a_3 t^3$$

$$\boldsymbol{r} = \begin{bmatrix} f(t_1) - y_1 \\ \vdots \\ f(t_i) - y_i \\ \vdots \\ f(t_n) - y_n \end{bmatrix} = \begin{bmatrix} a_o + a_1 t_1 + a_2 t_1^2 + a_3 t_1^3 - y_1 \\ \vdots \\ a_o + a_1 t_i + a_2 t_i^2 + a_3 t_i^3 - y_i \\ \vdots \\ a_o + a_1 t_n + a_2 t_n^2 + a_3 t_n^3 - y_n \end{bmatrix}$$

Least Squares Approximation

Consider n data points (t_i, y_i)

$$y = f(t) = a_o + a_1 t + a_2 t^2 + a_3 t^3$$

$$\boldsymbol{r} = \begin{bmatrix} f(t_1) - y_1 \\ \vdots \\ f(t_i) - y_i \\ \vdots \\ f(t_n) - y_n \end{bmatrix} = \begin{bmatrix} 1 & t_1 & t_1^2 & t_1^3 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & t_i & t_i^2 & t_i^3 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & t_n & t_n^2 & t_n^3 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} - \begin{bmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_n \end{bmatrix}$$

Least Squares Approximation

$$\begin{bmatrix} 1 & t_1 & t_1^2 & t_1^3 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & t_i & t_i^2 & t_i^3 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & t_n & t_n^2 & t_n^3 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_n \end{bmatrix} \qquad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

The parameters that result in a least squares approximation are the solution of:

$$\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{b}$$

Vandermonde Matrix

$$\begin{bmatrix} 1 & t_1 & t_1^2 & t_1^3 & \cdots & t_1^m \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & t_i & t_i^2 & t_i^3 & \cdots & t_i^m \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & t_n & t_n^2 & t_n^3 & \cdots & t_n^m \end{bmatrix}$$

The Vandermonde Matrix is well known to be ill-conditioned

Conditioning of the Normal Equations

$$\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{b}$$

$$\text{cond}\{\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A}\} \cong \text{cond}\{\boldsymbol{A}\}^{\mathbf{2}}$$

Can be very ill-conditioned.

Tikhonov Regularization

$$\begin{bmatrix} 1 & t_1 & t_1^2 & t_1^3 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & t_i & t_i^2 & t_i^3 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & t_n & t_n^2 & t_n^3 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_n \end{bmatrix}$$

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

- If the problem is ill-conditioned
- Impose an additional constraint that $\|\boldsymbol{x}\|$ is small.

Tikhonov Regularization

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

The solution of $\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{b}$

minimizes the residual: $\min\limits_{\boldsymbol{x}} ||\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}||_{\boldsymbol{2}}^{\boldsymbol{2}}$

Instead, minimize: $\min\limits_{\boldsymbol{x}} \left\{ ||\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}||_{\boldsymbol{2}}^{\boldsymbol{2}} + \lambda^2 ||\boldsymbol{x}||_2^2 \right\}$

Tikhonov Regularization

$$\begin{bmatrix} \boldsymbol{A} \\ \lambda \boldsymbol{U} \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} \boldsymbol{b} \\ \boldsymbol{0} \end{bmatrix}$$

$$\boldsymbol{r} = \begin{bmatrix} \boldsymbol{A} \\ \lambda \boldsymbol{U} \end{bmatrix} \boldsymbol{x} - \begin{bmatrix} \boldsymbol{b} \\ \boldsymbol{0} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b} \\ \lambda \boldsymbol{x} \end{bmatrix}$$

$$\|\boldsymbol{r}\|_2^2 = \boldsymbol{r}^T \boldsymbol{r} = \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2 + \lambda^2 \|\boldsymbol{x}\|_2^2$$

Tikhonov Regularization

$$\begin{bmatrix} \boldsymbol{A} \\ \lambda \boldsymbol{U} \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} \boldsymbol{b} \\ \boldsymbol{0} \end{bmatrix}$$

Normal Equation: $\begin{bmatrix} \boldsymbol{A}^T & \lambda \boldsymbol{U} \end{bmatrix} \begin{bmatrix} \boldsymbol{A} \\ \lambda \boldsymbol{U} \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} \boldsymbol{A}^T & \lambda \boldsymbol{U} \end{bmatrix} \begin{bmatrix} \boldsymbol{b} \\ \boldsymbol{0} \end{bmatrix}$

$$(\boldsymbol{A}^T \boldsymbol{A} + \lambda^2 \boldsymbol{U}) \boldsymbol{x} = \boldsymbol{A}^T \boldsymbol{b}$$

Conditioning of the Normal Equations

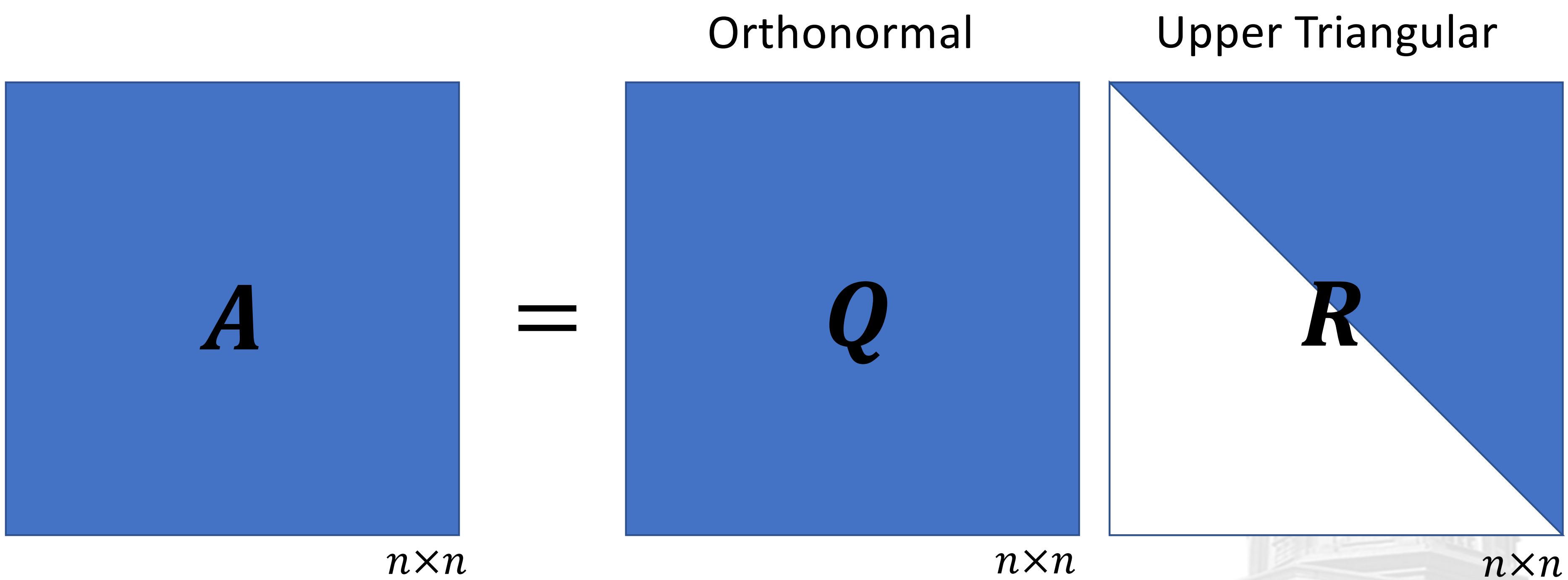
$$\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{b}$$

$$\text{cond}\{\boldsymbol{A}^T\boldsymbol{A}\} \cong \text{cond}\{\boldsymbol{A}\}^{\mathbf{2}}$$

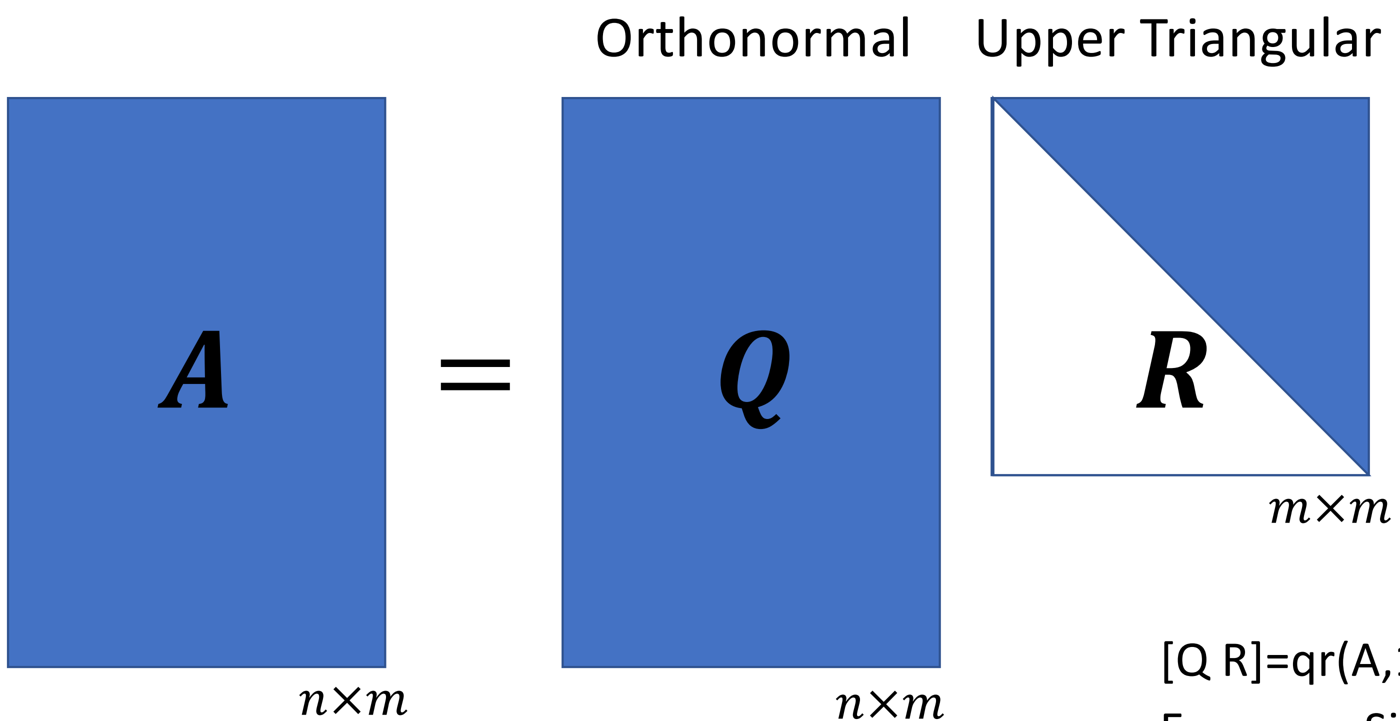
Can be very ill-conditioned.

Can we avoid the Normal Equations?

QR Decomposition



QR Decomposition

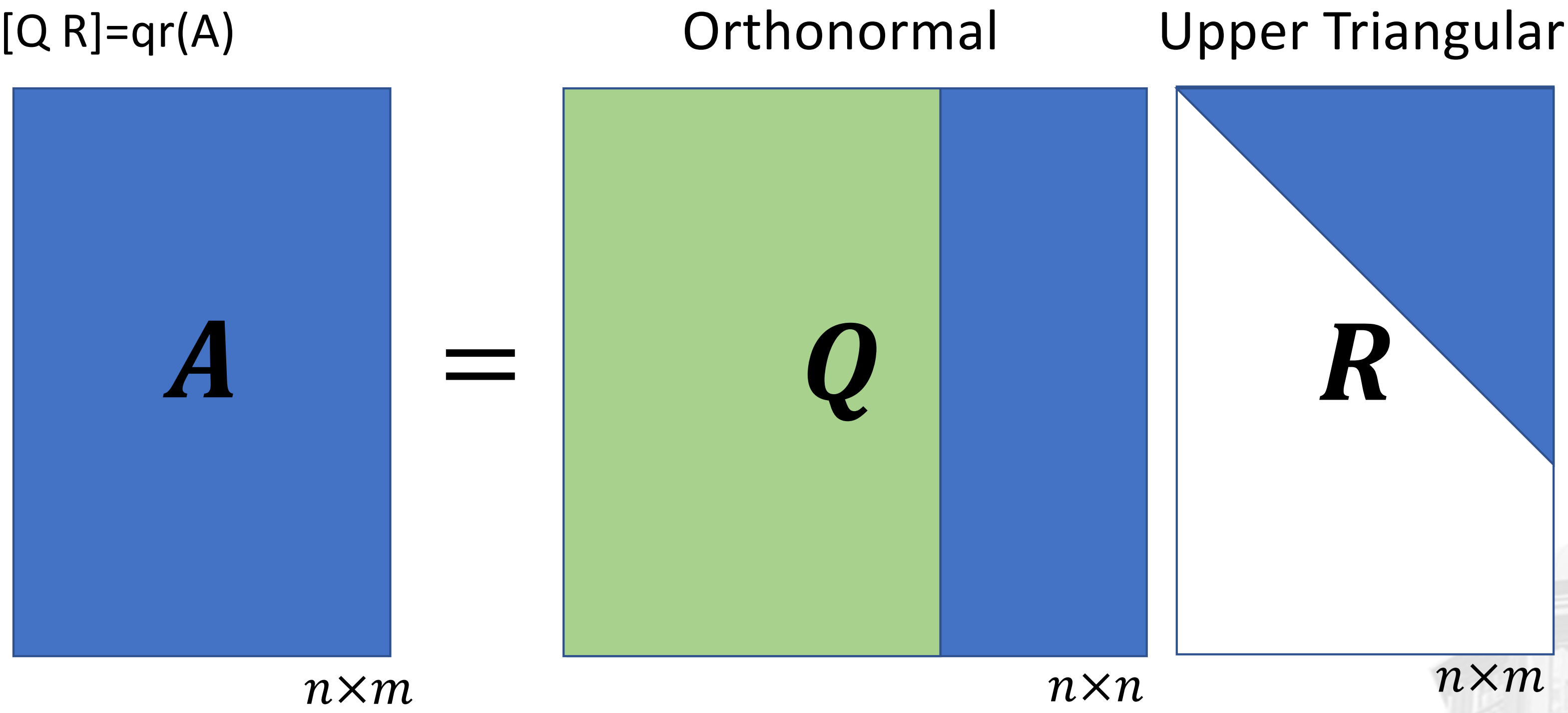


[Q R]=qr(A,1)

Economy-Size

QR Decomposition

[Q R]=qr(A)

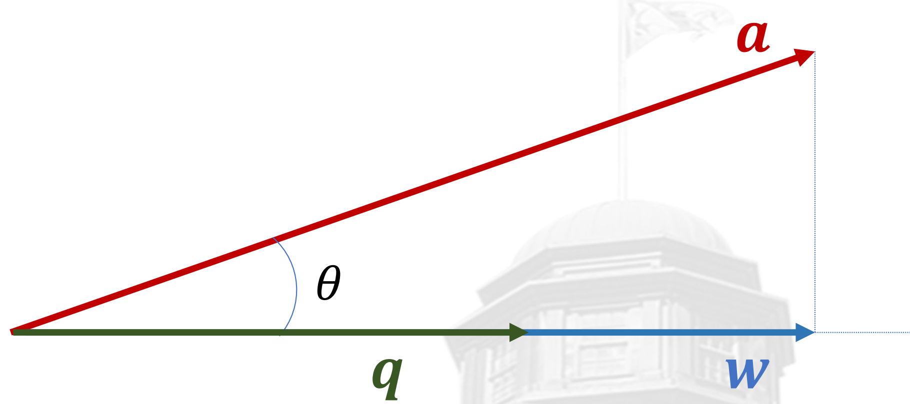


Projectors Review

$$\|\boldsymbol{q}\|_2 = \mathbf{1}$$

$$\langle \boldsymbol{a}, \boldsymbol{q} \rangle = \boldsymbol{q}^T \boldsymbol{a} = \|\boldsymbol{q}\|_2 \|\boldsymbol{a}\|_2 \cos(\theta) = \|\boldsymbol{a}\|_2 \cos(\theta)$$

$$\boldsymbol{w} = \boldsymbol{q} \langle \boldsymbol{a}, \boldsymbol{q} \rangle = \boldsymbol{q}\boldsymbol{q}^T \boldsymbol{a}$$



Gram-Schmidt Algorithm

Orthonormal

$$\left[\begin{bmatrix} \\ \boldsymbol{a}_1 \\ \\ \end{bmatrix} \begin{bmatrix} \\ \boldsymbol{a}_2 \\ \\ \end{bmatrix} \cdots \begin{bmatrix} \\ \boldsymbol{a}_m \\ \\ \end{bmatrix}\right] = \left[\begin{bmatrix} \\ \boldsymbol{q}_1 \\ \\ \end{bmatrix} \begin{bmatrix} \\ \boldsymbol{q}_2 \\ \\ \end{bmatrix} \cdots \begin{bmatrix} \\ \boldsymbol{q}_m \\ \\ \end{bmatrix}\right] \begin{bmatrix} r_{11} & r_{12} & r_{13} & \cdots & r_{1m} \\ & r_{22} & r_{23} & \cdots & r_{2m} \\ & & r_{33} & \cdots & r_{3m} \\ & & & \ddots & \vdots \\ & & & & r_{mm} \end{bmatrix}$$

$$\boldsymbol{q}_i^T \boldsymbol{q}_j = \begin{cases} 1\,; i = j \\ 0\,; i \neq j \end{cases}$$

Gram-Schmidt Algorithm

$$\begin{bmatrix} \begin{bmatrix} \\ \boldsymbol{a}_1 \\ \\ \end{bmatrix} & \begin{bmatrix} \\ \boldsymbol{a}_2 \\ \\ \end{bmatrix} & \cdots & \begin{bmatrix} \\ \boldsymbol{a}_m \\ \\ \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} \\ \boldsymbol{q}_1 \\ \\ \end{bmatrix} & \begin{bmatrix} \\ \boldsymbol{q}_2 \\ \\ \end{bmatrix} & \cdots & \begin{bmatrix} \\ \boldsymbol{q}_m \\ \\ \end{bmatrix} \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & \cdots & r_{1m} \\ & r_{22} & r_{23} & \cdots & r_{2m} \\ & & r_{33} & \cdots & r_{3m} \\ & & & \ddots & \vdots \\ & & & & r_{mm} \end{bmatrix}$$

$$\left.\begin{aligned} \boldsymbol{a}_1 &= \boldsymbol{q}_1 r_{11} \\ \|\boldsymbol{q}_1\| &= 1 \end{aligned}\right\} \quad \begin{aligned} r_{11} &= \|\boldsymbol{a}_1\| \\ \boldsymbol{q}_1 &= \frac{\boldsymbol{a}_1}{r_{11}} \end{aligned}$$

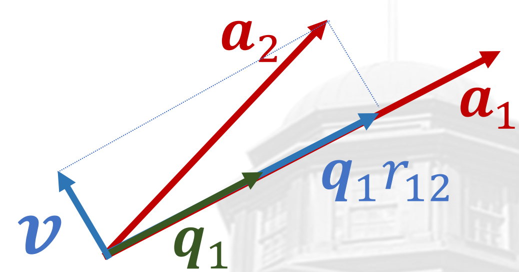
Gram-Schmidt Algorithm

$$\left[\begin{bmatrix} \\ \boldsymbol{a}_1 \\ \\ \end{bmatrix} \begin{bmatrix} \\ \boldsymbol{a}_2 \\ \\ \end{bmatrix} \cdots \begin{bmatrix} \\ \boldsymbol{a}_m \\ \\ \end{bmatrix}\right] = \left[\begin{bmatrix} \\ \boldsymbol{q}_1 \\ \\ \end{bmatrix} \begin{bmatrix} \\ \boldsymbol{q}_2 \\ \\ \end{bmatrix} \cdots \begin{bmatrix} \\ \boldsymbol{q}_m \\ \\ \end{bmatrix}\right] \begin{bmatrix} r_{11} & r_{12} & r_{13} & \cdots & r_{1m} \\ & r_{22} & r_{23} & \cdots & r_{2m} \\ & & r_{33} & \cdots & r_{3m} \\ & & & \ddots & \vdots \\ & & & & r_{mm} \end{bmatrix}$$

$$\boldsymbol{a}_2 = \boldsymbol{q}_1 r_{12} + \boldsymbol{q}_2 r_{22}$$

$$\boldsymbol{q}_1^T \boldsymbol{a}_2 = \boldsymbol{q}_1^T \boldsymbol{q}_1 r_{12} + \boldsymbol{q}_1^T \boldsymbol{q}_2 r_{22} = r_{12}$$

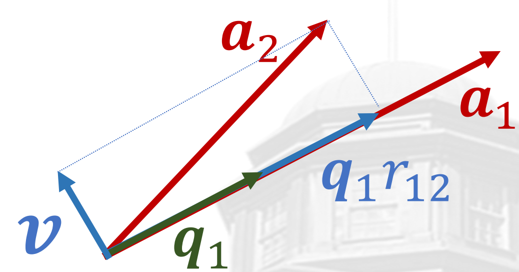
$$\boldsymbol{v} = \boldsymbol{a}_2 - \boldsymbol{q}_1 r_{12} = \boldsymbol{q}_2 r_{22}$$



Gram-Schmidt Algorithm

$$\left[\begin{bmatrix} \\ \mathbf{a}_1 \\ \\ \end{bmatrix} \begin{bmatrix} \\ \mathbf{a}_2 \\ \\ \end{bmatrix} \cdots \begin{bmatrix} \\ \mathbf{a}_m \\ \\ \end{bmatrix}\right] = \left[\begin{bmatrix} \\ \mathbf{q}_1 \\ \\ \end{bmatrix} \begin{bmatrix} \\ \mathbf{q}_2 \\ \\ \end{bmatrix} \cdots \begin{bmatrix} \\ \mathbf{q}_m \\ \\ \end{bmatrix}\right] \begin{bmatrix} r_{11} & r_{12} & r_{13} & \cdots & r_{1m} \\ & r_{22} & r_{23} & \cdots & r_{2m} \\ & & r_{33} & \cdots & r_{3m} \\ & & & \ddots & \vdots \\ & & & & r_{mm} \end{bmatrix}$$

$$\left.\begin{aligned} \mathbf{v} &= \mathbf{q}_2 r_{22} \\ \|\mathbf{q}_2\| &= 1 \end{aligned}\right\} \begin{aligned} r_{22} &= \|\mathbf{v}\| \\ \mathbf{q}_2 &= \frac{\mathbf{v}}{r_{22}} \end{aligned}$$



Gram-Schmidt Algorithm

$$\left[\begin{bmatrix} \\ \boldsymbol{a}_1 \\ \\ \end{bmatrix} \begin{bmatrix} \\ \boldsymbol{a}_2 \\ \\ \end{bmatrix} \cdots \begin{bmatrix} \\ \boldsymbol{a}_m \\ \\ \end{bmatrix}\right] = \left[\begin{bmatrix} \\ \boldsymbol{q}_1 \\ \\ \end{bmatrix} \begin{bmatrix} \\ \boldsymbol{q}_2 \\ \\ \end{bmatrix} \cdots \begin{bmatrix} \\ \boldsymbol{q}_m \\ \\ \end{bmatrix}\right] \begin{bmatrix} r_{11} & r_{12} & r_{13} & \cdots & r_{1m} \\ & r_{22} & r_{23} & \cdots & r_{2m} \\ & & r_{33} & \cdots & r_{3m} \\ & & & \ddots & \vdots \\ & & & & r_{mm} \end{bmatrix}$$

$$\boldsymbol{a}_3 = \boldsymbol{q}_1 r_{13} + \boldsymbol{q}_2 r_{23} + \boldsymbol{q}_3 r_{33}$$

$$\boldsymbol{q}_1^T \boldsymbol{a}_3 = r_{13}$$

$$\boldsymbol{q}_2^T \boldsymbol{a}_3 = r_{23}$$

$$\boldsymbol{v} = \boldsymbol{a}_3 - \boldsymbol{q}_1 r_{13} - \boldsymbol{q}_2 r_{23} = \boldsymbol{q}_3 r_{33}$$

Gram-Schmidt Algorithm

$$\begin{bmatrix} \begin{bmatrix} \\ \boldsymbol{a}_1 \\ \\ \end{bmatrix} & \begin{bmatrix} \\ \boldsymbol{a}_2 \\ \\ \end{bmatrix} & \cdots & \begin{bmatrix} \\ \boldsymbol{a}_m \\ \\ \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} \\ \boldsymbol{q}_1 \\ \\ \end{bmatrix} & \begin{bmatrix} \\ \boldsymbol{q}_2 \\ \\ \end{bmatrix} & \cdots & \begin{bmatrix} \\ \boldsymbol{q}_m \\ \\ \end{bmatrix} \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & \cdots & r_{1m} \\ & r_{22} & r_{23} & \cdots & r_{2m} \\ & & r_{33} & \cdots & r_{3m} \\ & & & \ddots & \vdots \\ & & & & r_{mm} \end{bmatrix}$$

$$\left.\begin{aligned} \boldsymbol{v} &= \boldsymbol{q}_3 r_{34} \\ \|\boldsymbol{q}_3\| &= 1 \end{aligned}\right\} \begin{aligned} r_{33} &= \|\boldsymbol{v}\| \\ \boldsymbol{q}_3 &= \frac{\boldsymbol{v}}{r_{33}} \end{aligned}$$