CENG 216 – Numerical Computation

Least Squares - Part I

Mustafa Özuysal
mustafaozuysal@iyte.edu.tr

April 3, 2022

İzmir Institute of Technology

Slide Credits

Slides are based on material from the main textbook:

"Numerical Analysis", The new international edition, 2ed, by Timothy Sauer

Least Squares and the Normal Equations

$\mathtt{A}\mathbf{x} = \mathbf{b}$ with no solutions

$$\begin{aligned} x_1 + x_2 &= 2 \\ x_1 - x_2 &= 1 \\ x_1 + x_2 &= 3 \end{aligned}$$

Inconsistent Systems of Equations

$\mathtt{A}\mathbf{x} = \mathbf{b}$ with no solutions $\rightarrow$ Equations are inconsistent.

$$
\begin{aligned}
x_1 + x_2 &= 2 \\
x_1 - x_2 &= 1 \\
x_1 + x_2 &= 3
\end{aligned}
$$

$\mathbf{Ax} = \mathbf{b}$ with no solutions $\rightarrow$ Equations are inconsistent.

$$\left.\begin{aligned} x_1 + x_2 &= 2 \\ x_1 - x_2 &= 1 \\ x_1 + x_2 &= 3 \end{aligned}\right\} \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

Inconsistent Systems of Equations

$\mathbf{Ax} = \mathbf{b}$ with no solutions $\rightarrow$ Equations are inconsistent.

$$\left.\begin{aligned} x_1 + x_2 &= 2 \\ x_1 - x_2 &= 1 \\ x_1 + x_2 &= 3 \end{aligned}\right\} \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

$$x_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix},$$

and there is no x_1, x_2 pair that satisfies this equality.

Inconsistent Systems of Equations: General Case

In general,

$$\mathtt{A}_{m\times n}\,\mathbf{x} = \begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_n\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ \vdots\\ x_n\end{bmatrix}$$

$$= x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \ldots + x_n\mathbf{v}_n$$

$$= \mathbf{b},$$

where $\mathbf{v}_i \in \mathbb{R}^m$ and $\mathbf{b} \in \mathbb{R}^m$.

In general,

$$\begin{aligned}
\mathtt{A}_{m\times n}\,\mathbf{x} &= \begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_n\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ \vdots\\ x_n\end{bmatrix}\\
&= x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \ldots + x_n\mathbf{v}_n\\
&= \mathbf{b},
\end{aligned}$$

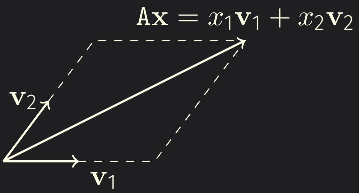
where $\mathbf{v}_i \in \mathbb{R}^m$ and $\mathbf{b} \in \mathbb{R}^m$. Whenever $m > n$, there is the possibility that the equations will be inconsistent.

Inconsistent Systems of Equations: Geometric View

$$\mathtt{A}\mathbf{x} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 = x_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} = \mathbf{b}$$

Inconsistent Systems of Equations: Geometric View

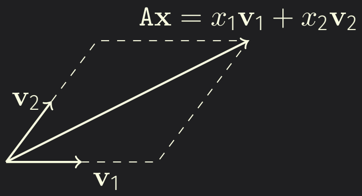
$$\mathtt{A}\mathbf{x} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 = x_1\begin{bmatrix}1\\1\\1\end{bmatrix} + x_2\begin{bmatrix}1\\-1\\1\end{bmatrix} = \begin{bmatrix}2\\1\\3\end{bmatrix} = \mathbf{b}$$



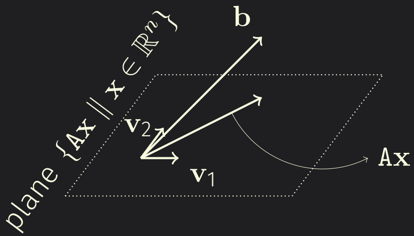
Linear combination of vectors creates another vector in the same plane.

Inconsistent Systems of Equations: Geometric View

$$\mathtt{A}\mathbf{x} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 = x_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} = \mathbf{b}$$

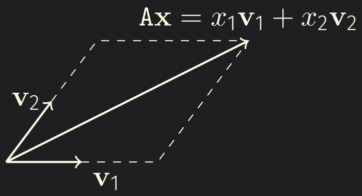


Linear combination of vectors creates another vector in the same plane.

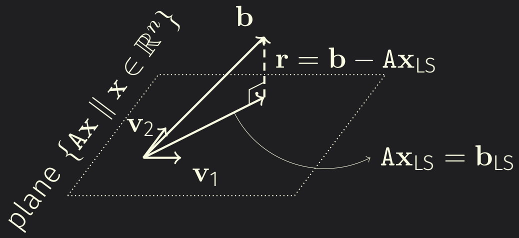


Inconsistent System

$$\mathtt{A}\mathbf{x} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 = x_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} = \mathbf{b}$$



Linear combination of vectors creates another vector in the same plane.



Inconsistent System

Def: Transpose of $\mathtt{A}$ is $\mathtt{A}^\top$ with rows as the columns of $\mathtt{A}$.

Def: Transpose of $\mathtt{A}$ is $\mathtt{A}^\top$ with rows as the columns of $\mathtt{A}$.
If $\mathtt{A}$ is $m \times n$ then $\mathtt{A}^\top$ is $n \times m$.

Def: Transpose of $\mathtt{A}$ is $\mathtt{A}^\top$ with rows as the columns of $\mathtt{A}$.
If $\mathtt{A}$ is $m \times n$ then $\mathtt{A}^\top$ is $n \times m$.

$$(\mathtt{A} + \mathtt{B})^\top = \mathtt{A}^\top + \mathtt{B}^\top \text{ and } (\mathtt{AB})^\top = \mathtt{B}^\top \mathtt{A}^\top.$$

Reminder: Vector and Matrix Transpose

Def: Transpose of $\mathtt{A}$ is $\mathtt{A}^\top$ with rows as the columns of $\mathtt{A}$.
If $\mathtt{A}$ is $m \times n$ then $\mathtt{A}^\top$ is $n \times m$.

$$(\mathtt{A} + \mathtt{B})^\top = \mathtt{A}^\top + \mathtt{B}^\top \text{ and } (\mathtt{AB})^\top = \mathtt{B}^\top \mathtt{A}^\top.$$

$$\mathbf{u}^\top \mathbf{v} = \mathbf{v}^\top \mathbf{u} = \|\mathbf{u}\| \, \|\mathbf{v}\| \cos \theta_{\mathbf{uv}},$$

where $\theta_{\mathbf{uv}}$ is the angle between $\mathbf{u}$ and $\mathbf{v}$.

Def: Transpose of $\mathtt{A}$ is $\mathtt{A}^\top$ with rows as the columns of $\mathtt{A}$.
If $\mathtt{A}$ is $m \times n$ then $\mathtt{A}^\top$ is $n \times m$.

$$(\mathtt{A} + \mathtt{B})^\top = \mathtt{A}^\top + \mathtt{B}^\top \text{ and } (\mathtt{A}\mathtt{B})^\top = \mathtt{B}^\top \mathtt{A}^\top.$$

$$\mathbf{u}^\top \mathbf{v} = \mathbf{v}^\top \mathbf{u} = \|\mathbf{u}\| \, \|\mathbf{v}\| \cos\theta_{\mathbf{uv}},$$

where $\theta_{\mathbf{uv}}$ is the angle between $\mathbf{u}$ and $\mathbf{v}$. If $\mathbf{u} \perp \mathbf{v}$ then
$\theta_{\mathbf{uv}} = 90° \implies \cos\theta_{\mathbf{uv}} = 0 \implies \mathbf{u}^\top \mathbf{v} = 0.$

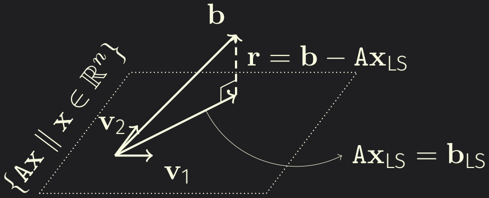
Def: Transpose of $\mathtt{A}$ is $\mathtt{A}^\top$ with rows as the columns of $\mathtt{A}$.
If $\mathtt{A}$ is $m \times n$ then $\mathtt{A}^\top$ is $n \times m$.

$$(\mathtt{A} + \mathtt{B})^\top = \mathtt{A}^\top + \mathtt{B}^\top \text{ and } (\mathtt{AB})^\top = \mathtt{B}^\top \mathtt{A}^\top.$$

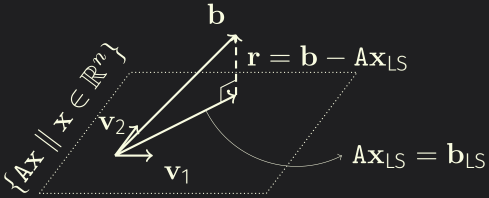
$$\mathbf{u}^\top \mathbf{v} = \mathbf{v}^\top \mathbf{u} = \|\mathbf{u}\| \, \|\mathbf{v}\| \cos \theta_{\mathbf{uv}},$$

where $\theta_{\mathbf{uv}}$ is the angle between $\mathbf{u}$ and $\mathbf{v}$. If $\mathbf{u} \perp \mathbf{v}$ then
$\theta_{\mathbf{uv}} = 90° \implies \cos \theta_{\mathbf{uv}} = 0 \implies \mathbf{u}^\top \mathbf{v} = 0.$

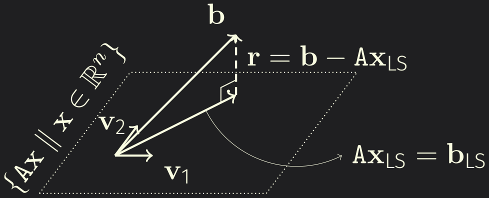
Normal Equations



$$(\mathbf{b} - \mathtt{A}\mathbf{x}_{\text{LS}}) \perp \{\mathtt{A}\mathbf{x} \parallel \mathbf{x} \in \mathbb{R}^n\}$$

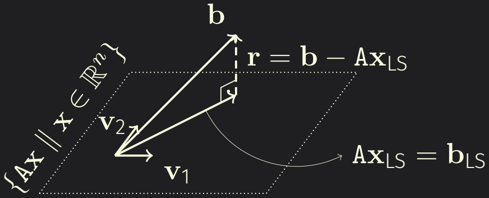


$$(\mathbf{b} - \mathtt{A}\mathbf{x}_{\mathsf{LS}}) \perp \{\mathtt{A}\mathbf{x} \parallel \mathbf{x} \in \mathbb{R}^n\} \implies (\mathtt{A}\mathbf{x})^\top (\mathbf{b} - \mathtt{A}\mathbf{x}_{\mathsf{LS}}) = 0,\ \forall \mathbf{x} \in \mathbb{R}^n$$

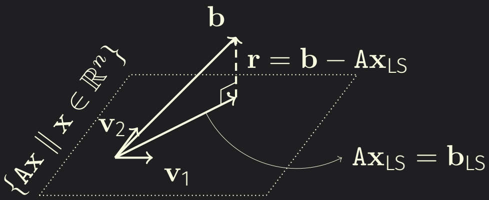


$$
\begin{aligned}
(\mathbf{b} - \mathtt{A}\mathbf{x}_{\mathsf{LS}}) \perp \{\mathtt{A}\mathbf{x} \parallel \mathbf{x} \in \mathbb{R}^n\} &\implies (\mathtt{A}\mathbf{x})^\top (\mathbf{b} - \mathtt{A}\mathbf{x}_{\mathsf{LS}}) = 0,\ \forall \mathbf{x} \in \mathbb{R}^n \\
&\implies \mathbf{x}^\top \mathtt{A}^\top (\mathbf{b} - \mathtt{A}\mathbf{x}_{\mathsf{LS}}) = 0,\ \forall \mathbf{x} \in \mathbb{R}^n
\end{aligned}
$$

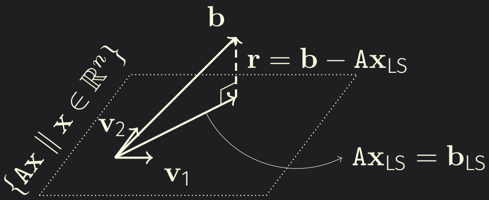
Normal Equations



$$
\begin{aligned}
(\mathbf{b} - \mathtt{A}\mathbf{x}_{\mathsf{LS}}) \perp \{\mathtt{A}\mathbf{x} \parallel \mathbf{x} \in \mathbb{R}^n\} \implies & (\mathtt{A}\mathbf{x})^\top (\mathbf{b} - \mathtt{A}\mathbf{x}_{\mathsf{LS}}) = 0, \forall \mathbf{x} \in \mathbb{R}^n \\
\implies & \mathbf{x}^\top \mathtt{A}^\top (\mathbf{b} - \mathtt{A}\mathbf{x}_{\mathsf{LS}}) = 0, \forall \mathbf{x} \in \mathbb{R}^n \\
\implies & \mathtt{A}^\top (\mathbf{b} - \mathtt{A}\mathbf{x}_{\mathsf{LS}}) = 0
\end{aligned}
$$

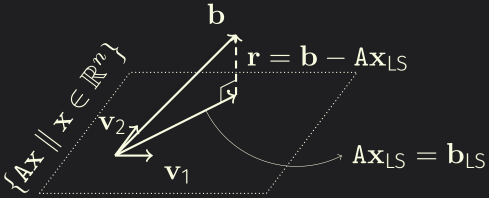


$$
\begin{aligned}
(\mathbf{b} - \mathtt{A}\mathbf{x}_{\mathsf{LS}}) \perp \{\mathtt{A}\mathbf{x} \,\|\, \mathbf{x} \in \mathbb{R}^n\} &\implies (\mathtt{A}\mathbf{x})^\top (\mathbf{b} - \mathtt{A}\mathbf{x}_{\mathsf{LS}}) = 0,\ \forall \mathbf{x} \in \mathbb{R}^n \\
&\implies \mathbf{x}^\top \mathtt{A}^\top (\mathbf{b} - \mathtt{A}\mathbf{x}_{\mathsf{LS}}) = 0,\ \forall \mathbf{x} \in \mathbb{R}^n \\
&\implies \mathtt{A}^\top (\mathbf{b} - \mathtt{A}\mathbf{x}_{\mathsf{LS}}) = 0 \\
&\implies \mathtt{A}^\top \mathtt{A}\mathbf{x}_{\mathsf{LS}} = \mathtt{A}^\top \mathbf{b} \quad \text{(Normal Equations)}
\end{aligned}
$$



$$
\begin{aligned}
(\mathbf{b} - \mathtt{A}\mathbf{x}_{\mathsf{LS}}) \perp \{\mathtt{A}\mathbf{x} \mid\mid \mathbf{x} \in \mathbb{R}^n\} \quad &\Longrightarrow \quad (\mathtt{A}\mathbf{x})^\top (\mathbf{b} - \mathtt{A}\mathbf{x}_{\mathsf{LS}}) = 0,\ \forall \mathbf{x} \in \mathbb{R}^n \\
&\Longrightarrow \quad \mathbf{x}^\top \mathtt{A}^\top (\mathbf{b} - \mathtt{A}\mathbf{x}_{\mathsf{LS}}) = 0,\ \forall \mathbf{x} \in \mathbb{R}^n \\
&\Longrightarrow \quad \mathtt{A}^\top (\mathbf{b} - \mathtt{A}\mathbf{x}_{\mathsf{LS}}) = 0 \\
&\Longrightarrow \quad \mathtt{A}^\top \mathtt{A}\mathbf{x}_{\mathsf{LS}} = \mathtt{A}^\top \mathbf{b} \quad \text{(Normal Equations)}
\end{aligned}
$$

$\mathbf{r}$ is called the residual vector.



$$
\begin{aligned}
(\mathbf{b} - \mathtt{A}\mathbf{x}_{\mathsf{LS}}) \perp \{\mathtt{A}\mathbf{x} \parallel \mathbf{x} \in \mathbb{R}^n\} \quad &\implies (\mathtt{A}\mathbf{x})^\top (\mathbf{b} - \mathtt{A}\mathbf{x}_{\mathsf{LS}}) = 0, \ \forall \mathbf{x} \in \mathbb{R}^n \\
&\implies \mathbf{x}^\top \mathtt{A}^\top (\mathbf{b} - \mathtt{A}\mathbf{x}_{\mathsf{LS}}) = 0, \ \forall \mathbf{x} \in \mathbb{R}^n \\
&\implies \mathtt{A}^\top (\mathbf{b} - \mathtt{A}\mathbf{x}_{\mathsf{LS}}) = 0 \\
&\implies \mathtt{A}^\top \mathtt{A}\mathbf{x}_{\mathsf{LS}} = \mathtt{A}^\top \mathbf{b} \quad \text{(Normal Equations)}
\end{aligned}
$$

$\mathbf{r}$ is called the residual vector. The squared error (SE) and the root mean squared error (RMSE) are defined to be

$$
\mathsf{SE} = \|\mathbf{r}\|^2 = r_1^2 + r_2^2 + \ldots + r_n^2 \text{ and } \mathsf{RMSE} = \sqrt{\frac{\mathsf{SE}}{m}} = \frac{\|\mathbf{r}\|}{\sqrt{m}}
$$

Example

Find the least squares solution $\mathbf{x}_{\text{LS}}$ for $\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$.

Example

Find the least squares solution $\mathbf{x}_{\text{LS}}$ for $\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$.

$$\mathtt{A}^\top \mathtt{A} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$$

Example

Find the least squares solution $\mathbf{x}_{\text{LS}}$ for $\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$.

$$\mathtt{A}^\top \mathtt{A} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$$

$$\mathtt{A}^\top \mathbf{b} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

Example

Find the least squares solution $\mathbf{x}_{\text{LS}}$ for $\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$.

$$\mathtt{A}^\top \mathtt{A} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$$

$$\mathtt{A}^\top \mathbf{b} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

Solving the system of normal equations $\mathtt{A}^\top \mathtt{A} \mathbf{x}_{\text{LS}} = \mathtt{A}^\top \mathbf{b}$

$$\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix} \implies \mathbf{x}_{\text{LS}} = \begin{bmatrix} \frac{7}{4} \\ \frac{3}{4} \end{bmatrix}.$$

Find the least squares solution $\mathbf{x}_{\text{LS}}$ for $\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$.

$$\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix} \implies \mathbf{x}_{\text{LS}} = \begin{bmatrix} \frac{7}{4} \\ \frac{3}{4} \end{bmatrix}.$$

Find the least squares solution $\mathbf{x}_{\text{LS}}$ for $\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$.

$$\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix} \implies \mathbf{x}_{\text{LS}} = \begin{bmatrix} \frac{7}{4} \\ \frac{3}{4} \end{bmatrix}.$$

Then,

$$\mathbf{r} = \mathbf{b} - \mathtt{A}\mathbf{x}_{\text{LS}} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} - \begin{bmatrix} 2.5 \\ 1 \\ 2.5 \end{bmatrix} = \begin{bmatrix} -0.5 \\ 0.0 \\ 0.5 \end{bmatrix}$$

Find the least squares solution $\mathbf{x}_{\text{LS}}$ for $\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$.

$$\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix} \implies \mathbf{x}_{\text{LS}} = \begin{bmatrix} \frac{7}{4} \\ \frac{3}{4} \end{bmatrix}.$$

Then,

$$\mathbf{r} = \mathbf{b} - \mathtt{A}\mathbf{x}_{\text{LS}} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} - \begin{bmatrix} 2.5 \\ 1 \\ 2.5 \end{bmatrix} = \begin{bmatrix} -0.5 \\ 0.0 \\ 0.5 \end{bmatrix}$$

$$\text{SE} = (-0.5)^2 + (0)^2 + (0.5)^2 = 0.5$$

Find the least squares solution $\mathbf{x}_{\text{LS}}$ for $\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$.

$$\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix} \implies \mathbf{x}_{\text{LS}} = \begin{bmatrix} \frac{7}{4} \\ \frac{3}{4} \end{bmatrix}.$$

Then,

$$\mathbf{r} = \mathbf{b} - \mathtt{A}\mathbf{x}_{\text{LS}} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} - \begin{bmatrix} 2.5 \\ 1 \\ 2.5 \end{bmatrix} = \begin{bmatrix} -0.5 \\ 0.0 \\ 0.5 \end{bmatrix}$$

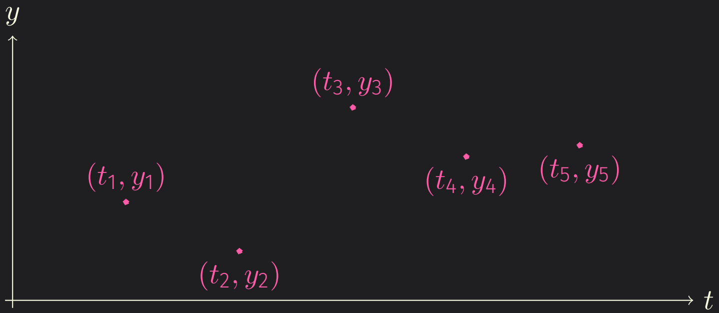
$$\text{SE} = (-0.5)^2 + (0)^2 + (0.5)^2 = 0.5$$

$$\text{RMSE} = \sqrt{\frac{\text{SE}}{3}} \approx 0.408$$

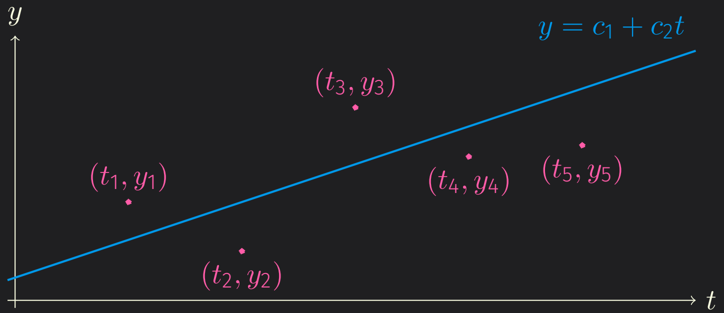
Model Fitting

Fitting Models to Data

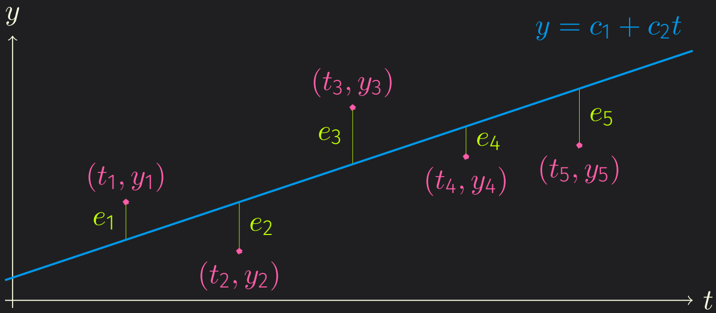
- Let $(t_1, y_1), (t_2, y_2), \ldots, (t_n, y_n)$ be a set of data points.



- Let $(t_1, y_1), (t_2, y_2), \ldots, (t_n, y_n)$ be a set of data points.
- Given a parametric model, such as the line $y = c_1 + c_2 t$ with parameters c_1 and c_2, find the "best" model that "fits" to the data points.

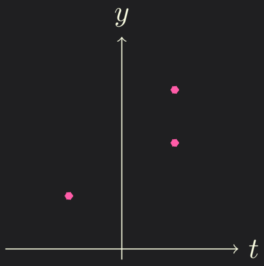


- Let $(t_1, y_1), (t_2, y_2), \ldots, (t_n, y_n)$ be a set of data points.
- Given a parametric model, such as the line $y = c_1 + c_2 t$ with parameters c_1 and c_2, find the "best" model that "fits" to the data points.
- The best parameters minimize the model error $\sum_{i=0}^{n} e_i^2$.



Find the line $y = c_1 + c_2 t$ that best fits the three data points $(t, y) = (1, 2), (-1, 1),$ and $(1, 3)$.

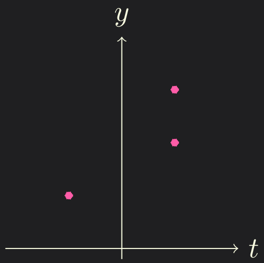
$$\begin{aligned} c_1 + c_2(1) &= 2 \\ c_1 + c_2(-1) &= 1 \\ c_1 + c_2(1) &= 3 \end{aligned}$$



Example

Find the line $y = c_1 + c_2 t$ that best fits the three data points $(t, y) = (1, 2), (-1, 1)$, and $(1, 3)$.

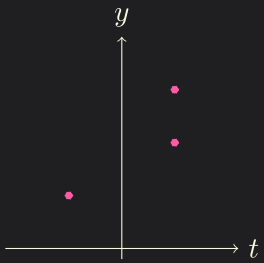
$$\left.\begin{aligned} c_1 + c_2(1) &= 2 \\ c_1 + c_2(-1) &= 1 \\ c_1 + c_2(1) &= 3 \end{aligned}\right\} \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$



Example

Find the line $y = c_1 + c_2 t$ that best fits the three data points $(t, y) = (1, 2), (-1, 1),$ and $(1, 3)$.

$$\left.\begin{aligned} c_1 + c_2(1) &= 2 \\ c_1 + c_2(-1) &= 1 \\ c_1 + c_2(1) &= 3 \end{aligned}\right\} \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} \to \begin{aligned} c_1 &= \frac{7}{4} \\ c_2 &= \frac{3}{4} \end{aligned}$$

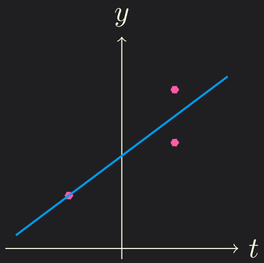


Find the line $y = c_1 + c_2 t$ that best fits the three data points $(t, y) = (1, 2), (-1, 1),$ and $(1, 3)$.

$$\left.\begin{aligned} c_1 + c_2(1) &= 2 \\ c_1 + c_2(-1) &= 1 \\ c_1 + c_2(1) &= 3 \end{aligned}\right\} \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} \rightarrow \begin{aligned} c_1 &= \frac{7}{4} \\ c_2 &= \frac{3}{4} \end{aligned}$$

So the best fitting line is

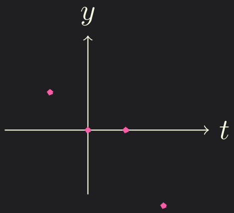
$$y = \frac{7}{4} + \frac{3}{4}t$$



Example

Find the line $y = c_1 + c_2 t$ and the parabola $y = c_1 + c_2 t + c_3 t^2$ that best fits the four data points $(t, y) = (-1, 1), (0, 0), (1, 0)$, and $(2, -2)$.

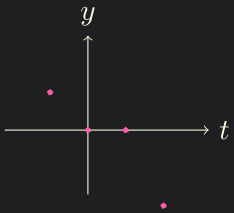
$$\begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -2 \end{bmatrix}$$



Example

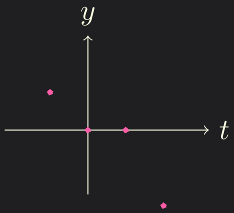
Find the line $y = c_1 + c_2 t$ and the parabola $y = c_1 + c_2 t + c_3 t^2$ that best fits the four data points $(t, y) = (-1, 1), (0, 0), (1, 0),$ and $(2, -2)$.

$$\left.\begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -2 \end{bmatrix}\right\} \begin{bmatrix} 4 & 2 \\ 2 & 6 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} -1 \\ -5 \end{bmatrix}$$



Find the line $y = c_1 + c_2 t$ and the parabola $y = c_1 + c_2 t + c_3 t^2$ that best fits the four data points $(t, y) = (-1, 1), (0, 0), (1, 0),$ and $(2, -2)$.

$$\left.\begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -2 \end{bmatrix} \right\} \begin{bmatrix} 4 & 2 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} -1 \\ -5 \end{bmatrix} \rightarrow \begin{matrix} c_1 = 0.2 \\ c_2 = -0.9 \end{matrix}$$



Example

Find the line $y = c_1 + c_2 t$ and the parabola $y = c_1 + c_2 t + c_3 t^2$ that best fits the four data points $(t, y) = (-1, 1), (0, 0), (1, 0),$ and $(2, -2)$.

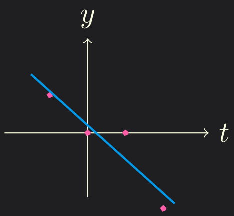
$$\left.\begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -2 \end{bmatrix} \right\} \begin{bmatrix} 4 & 2 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} -1 \\ -5 \end{bmatrix} \rightarrow \begin{array}{l} c_1 = 0.2 \\ c_2 = -0.9 \end{array}$$

So the best fitting line is

$$y = 0.2 - 0.9t$$

$$\text{SE} = 0.7$$

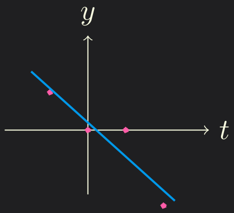
$$\text{RMSE} = 0.418$$



Example

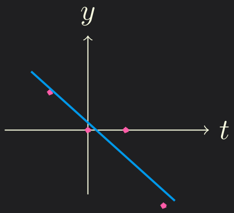
Find the line $y = c_1 + c_2 t$ and the parabola $y = c_1 + c_2 t + c_3 t^2$ that best fits the four data points $(t, y) = (-1, 1), (0, 0), (1, 0),$ and $(2, -2)$.

$$\begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -2 \end{bmatrix}$$



Find the line $y = c_1 + c_2 t$ and the parabola $y = c_1 + c_2 t + c_3 t^2$ that best fits the four data points $(t, y) = (-1, 1), (0, 0), (1, 0),$ and $(2, -2)$.

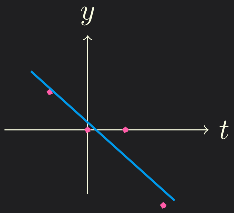
$$\left.\begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -2 \end{bmatrix} \right\} \begin{bmatrix} 4 & 2 & 6 \\ 2 & 6 & 8 \\ 6 & 8 & 18 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} -1 \\ -5 \\ -7 \end{bmatrix}$$



Example

Find the line $y = c_1 + c_2 t$ and the parabola $y = c_1 + c_2 t + c_3 t^2$ that best fits the four data points
$(t, y) = (-1, 1), (0, 0), (1, 0),$ and $(2, -2)$.

$$\left.\begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -2 \end{bmatrix} \right\} \begin{bmatrix} 4 & 2 & 6 \\ 2 & 6 & 8 \\ 6 & 8 & 18 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} -1 \\ -5 \\ -7 \end{bmatrix} \rightarrow \begin{aligned} c_1 &= 0.45 \\ c_2 &= -0.65 \\ c_3 &= -0.25 \end{aligned}$$



Example

Find the line $y = c_1 + c_2 t$ and the parabola $y = c_1 + c_2 t + c_3 t^2$ that best fits the four data points $(t, y) = (-1, 1), (0, 0), (1, 0),$ and $(2, -2)$.

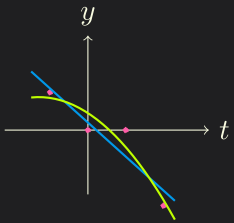
$$\left.\begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -2 \end{bmatrix} \right\} \begin{bmatrix} 4 & 2 & 6 \\ 2 & 6 & 8 \\ 6 & 8 & 18 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} -1 \\ -5 \\ -7 \end{bmatrix} \to \begin{matrix} c_1 = 0.45 \\ c_2 = -0.65 \\ c_3 = -0.25 \end{matrix}$$

So the best fitting parabola is

$$y = 0.45 - 0.65t - 0.25t^2$$

$$\text{SE} = 0.45$$

$$\text{RMSE} \approx 0.335$$



So far, to solve the least square problem $\mathbf{A}\mathbf{x} = \mathbf{b}$, we have solved the normal equations

$$\mathbf{A}^\top \mathbf{A}\mathbf{x} = \mathbf{A}^\top \mathbf{b}.$$

So far, to solve the least square problem $\mathtt{A}\mathbf{x} = \mathbf{b}$, we have solved the normal equations

$$\mathtt{A}^\top \mathtt{A}\mathbf{x} = \mathtt{A}^\top \mathbf{b}.$$

However, the numerically this is problematic since

$$\mathrm{cond}(\mathtt{A}^\top \mathtt{A}) = (\mathrm{cond}(\mathtt{A}))^2\,,$$

so if the condition number of $\mathtt{A}$ is large, the condition number of $\mathtt{A}^\top \mathtt{A}$ is even larger! We need a better approach that does not require explicitly forming $\mathtt{A}^\top \mathtt{A}$.

So far, to solve the least square problem $\mathtt{A}\mathbf{x} = \mathbf{b}$, we have solved the normal equations

$$\mathtt{A}^\top \mathtt{A}\mathbf{x} = \mathtt{A}^\top \mathbf{b}.$$

However, the numerically this is problematic since

$$\mathrm{cond}(\mathtt{A}^\top \mathtt{A}) = (\mathrm{cond}(\mathtt{A}))^2 ,$$

so if the condition number of $\mathtt{A}$ is large, the condition number of $\mathtt{A}^\top \mathtt{A}$ is even larger! We need a better approach that does not require explicitly forming $\mathtt{A}^\top \mathtt{A}$. $\rightarrow$QR Factorization

QR Factorization

- Given an input set of m-dimensional vectors, find an **orthogonal** coordinate system for the subspace spanned by the set (all vectors that can be reached by linear combination).

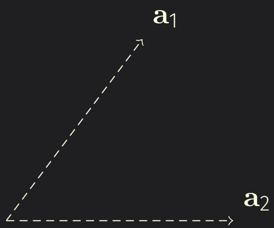
Gram-Schmidt Orthogonalization

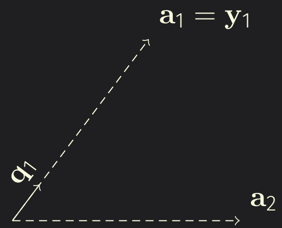
- Given an input set of m-dimensional vectors, find an orthogonal coordinate system for the subspace spanned by the set (all vectors that can be reached by linear combination).
- Input: n-linearly independent vectors $\{\mathbf{a}_i\}$.

- Given an input set of m-dimensional vectors, find an orthogonal coordinate system for the subspace spanned by the set (all vectors that can be reached by linear combination).
- **Input:** n-linearly independent vectors $\{\mathbf{a}_i\}$.
- **Output:** n mutually perpendicular unit vectors $\{\mathbf{q}_i\}$ that forms an orthogonal coordinate system for $\mathrm{span}\{\mathbf{a}_i\}$

$$\|\mathbf{q}_i\| = 1 \implies \mathbf{q}_i^\top \mathbf{q}_i = 1, \quad \forall i = 1, \ldots, n.$$
$$\mathbf{q}_i \perp \mathbf{q}_j \implies \mathbf{q}_i^\top \mathbf{q}_j = 0, \quad \forall i, j \in \{1, \ldots, n\} \text{ and } i \neq j.$$

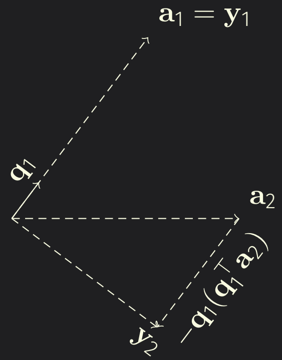
Let $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n \in \mathbb{R}^m$ be linearly independent ($m \geq n$).





Let $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n \in \mathbb{R}^m$ be linearly independent ($m \geq n$).

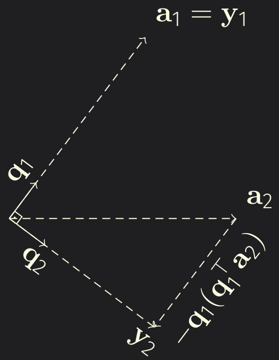
$\mathbf{y}_1 = \mathbf{a}_1,$ $\qquad$ $\mathbf{q}_1 = \dfrac{\mathbf{y}_1}{\|\mathbf{y}_1\|}$



Let $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n \in \mathbb{R}^m$ be linearly independent ($m \geq n$).

$$\mathbf{y}_1 = \mathbf{a}_1, \qquad \mathbf{q}_1 = \frac{\mathbf{y}_1}{\|\mathbf{y}_1\|}$$

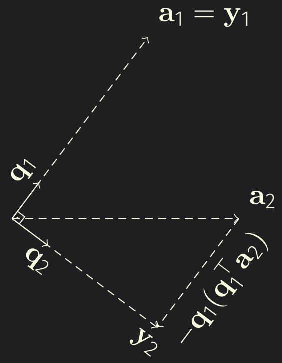
$$\mathbf{y}_2 = \mathbf{a}_2 - \mathbf{q}_1(\mathbf{q}_1^\top \mathbf{a}_2)$$



Let $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n \in \mathbb{R}^m$ be linearly independent ($m \geq n$).

$$\mathbf{y}_1 = \mathbf{a}_1, \qquad \mathbf{q}_1 = \frac{\mathbf{y}_1}{\|\mathbf{y}_1\|}$$

$$\mathbf{y}_2 = \mathbf{a}_2 - \mathbf{q}_1(\mathbf{q}_1^\top \mathbf{a}_2), \qquad \mathbf{q}_2 = \frac{\mathbf{y}_2}{\|\mathbf{y}_2\|}$$



Let $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n \in \mathbb{R}^m$ be linearly independent ($m \geq n$).

$$\mathbf{y}_1 = \mathbf{a}_1, \qquad \mathbf{q}_1 = \frac{\mathbf{y}_1}{\|\mathbf{y}_1\|}$$

$$\mathbf{y}_2 = \mathbf{a}_2 - \mathbf{q}_1(\mathbf{q}_1^\top \mathbf{a}_2), \qquad \mathbf{q}_2 = \frac{\mathbf{y}_2}{\|\mathbf{y}_2\|}$$

$$\vdots \qquad\qquad \vdots$$

$$\begin{aligned}\mathbf{y}_j = \mathbf{a}_j &- \mathbf{q}_1(\mathbf{q}_1^\top \mathbf{a}_j) - \mathbf{q}_2(\mathbf{q}_2^\top \mathbf{a}_j) \\ &- \ldots - \mathbf{q}_{j-1}(\mathbf{q}_{j-1}^\top \mathbf{a}_j), \qquad \mathbf{q}_j = \frac{\mathbf{y}_j}{\|\mathbf{y}_j\|}\end{aligned}$$

Gram-Schmidt Orthogonalization

Prove that $\mathbf{q}_i \perp \mathbf{y}_j, \quad \forall i < j$: Proof-by-induction

Gram-Schmidt Orthogonalization

Prove that $\mathbf{q}_i \perp \mathbf{y}_j, \quad \forall i < j$: Proof-by-induction

- Prove that $\mathbf{q}_1 \perp \mathbf{y}_2$ hence $\mathbf{q}_1 \perp \mathbf{q}_2$ (Exercise)

Prove that $\mathbf{q}_i \perp \mathbf{y}_j, \quad \forall i < j$: Proof-by-induction

- Prove that $\mathbf{q}_1 \perp \mathbf{y}_2$ hence $\mathbf{q}_1 \perp \mathbf{q}_2$ (Exercise)
- Assume $\mathbf{q}_i \perp \mathbf{q}_k, \quad \forall k \leq j - 1, k \neq i$,
  Prove $\mathbf{q}_i \perp \mathbf{y}_j$:

Prove that $\mathbf{q}_i \perp \mathbf{y}_j, \quad \forall i < j$: Proof-by-induction

- Prove that $\mathbf{q}_1 \perp \mathbf{y}_2$ hence $\mathbf{q}_1 \perp \mathbf{q}_2$ (Exercise)
- Assume $\mathbf{q}_i \perp \mathbf{q}_k, \quad \forall k \leq j-1, k \neq i$,
  Prove $\mathbf{q}_i \perp \mathbf{y}_j$:

$$\mathbf{q}_i^\top \mathbf{y}_j = \mathbf{q}_i^\top \left(\mathbf{a}_j - \mathbf{q}_1(\mathbf{q}_1^\top \mathbf{a}_j) - \mathbf{q}_2(\mathbf{q}_2^\top \mathbf{a}_j) - \ldots - \mathbf{q}_{j-1}(\mathbf{q}_{j-1}^\top \mathbf{a}_j)\right)$$

Prove that $\mathbf{q}_i \perp \mathbf{y}_j, \quad \forall i < j$: Proof-by-induction

- Prove that $\mathbf{q}_1 \perp \mathbf{y}_2$ hence $\mathbf{q}_1 \perp \mathbf{q}_2$ (Exercise)
- Assume $\mathbf{q}_i \perp \mathbf{q}_k, \quad \forall k \leq j-1, k \neq i$,
  Prove $\mathbf{q}_i \perp \mathbf{y}_j$:

$$
\begin{aligned}
\mathbf{q}_i^\top \mathbf{y}_j &= \mathbf{q}_i^\top \left(\mathbf{a}_j - \mathbf{q}_1(\mathbf{q}_1^\top \mathbf{a}_j) - \mathbf{q}_2(\mathbf{q}_2^\top \mathbf{a}_j) - \ldots - \mathbf{q}_{j-1}(\mathbf{q}_{j-1}^\top \mathbf{a}_j)\right) \\
&= \mathbf{q}_i^\top \mathbf{a}_j - \underbrace{\mathbf{q}_i^\top \mathbf{q}_1}_{0}(\mathbf{q}_1^\top \mathbf{a}_j) - \ldots - \mathbf{q}_i^\top \mathbf{q}_i(\mathbf{q}_2^\top \mathbf{a}_j) - \ldots - \underbrace{\mathbf{q}_i^\top \mathbf{q}_{j-1}}_{0}(\mathbf{q}_{j-1}^\top \mathbf{a}_j)
\end{aligned}
$$

Prove that $\mathbf{q}_i \perp \mathbf{y}_j, \quad \forall i < j$: Proof-by-induction

- Prove that $\mathbf{q}_1 \perp \mathbf{y}_2$ hence $\mathbf{q}_1 \perp \mathbf{q}_2$ (Exercise)
- Assume $\mathbf{q}_i \perp \mathbf{q}_k, \quad \forall k \leq j-1, k \neq i$,
  Prove $\mathbf{q}_i \perp \mathbf{y}_j$:

$$
\begin{aligned}
\mathbf{q}_i^\top \mathbf{y}_j &= \mathbf{q}_i^\top \left(\mathbf{a}_j - \mathbf{q}_1(\mathbf{q}_1^\top \mathbf{a}_j) - \mathbf{q}_2(\mathbf{q}_2^\top \mathbf{a}_j) - \ldots - \mathbf{q}_{j-1}(\mathbf{q}_{j-1}^\top \mathbf{a}_j)\right) \\
&= \mathbf{q}_i^\top \mathbf{a}_j - \underbrace{\mathbf{q}_i^\top \mathbf{q}_1}_{0}(\mathbf{q}_1^\top \mathbf{a}_j) - \ldots - \mathbf{q}_i^\top \mathbf{q}_i(\mathbf{q}_2^\top \mathbf{a}_j) - \ldots - \underbrace{\mathbf{q}_i^\top \mathbf{q}_{j-1}}_{0}(\mathbf{q}_{j-1}^\top \mathbf{a}_j) \\
&= \mathbf{q}_i^\top \mathbf{a}_j - \underbrace{\mathbf{q}_i^\top \mathbf{q}_i}_{1}(\mathbf{q}_i^\top \mathbf{a}_j)
\end{aligned}
$$

Prove that $\mathbf{q}_i \perp \mathbf{y}_j, \quad \forall i < j$: Proof-by-induction

- Prove that $\mathbf{q}_1 \perp \mathbf{y}_2$ hence $\mathbf{q}_1 \perp \mathbf{q}_2$ (Exercise)
- Assume $\mathbf{q}_i \perp \mathbf{q}_k, \quad \forall k \leq j-1, k \neq i$,
  Prove $\mathbf{q}_i \perp \mathbf{y}_j$:

$$\begin{aligned}
\mathbf{q}_i^\top \mathbf{y}_j &= \mathbf{q}_i^\top \left(\mathbf{a}_j - \mathbf{q}_1(\mathbf{q}_1^\top \mathbf{a}_j) - \mathbf{q}_2(\mathbf{q}_2^\top \mathbf{a}_j) - \ldots - \mathbf{q}_{j-1}(\mathbf{q}_{j-1}^\top \mathbf{a}_j)\right) \\
&= \mathbf{q}_i^\top \mathbf{a}_j - \underbrace{\mathbf{q}_i^\top \mathbf{q}_1}_{0}(\mathbf{q}_1^\top \mathbf{a}_j) - \ldots - \mathbf{q}_i^\top \mathbf{q}_i(\mathbf{q}_2^\top \mathbf{a}_j) - \ldots - \underbrace{\mathbf{q}_i^\top \mathbf{q}_{j-1}}_{0}(\mathbf{q}_{j-1}^\top \mathbf{a}_j) \\
&= \mathbf{q}_i^\top \mathbf{a}_j - \underbrace{\mathbf{q}_i^\top \mathbf{q}_i}_{1}(\mathbf{q}_i^\top \mathbf{a}_j) \\
&= \mathbf{q}_i^\top \mathbf{a}_j - \mathbf{q}_i^\top \mathbf{a}_j \\
&= 0
\end{aligned}$$

Define $r_{jj} = \|\mathbf{y}_j\|$ and $r_{ij} = \mathbf{q}_i^\top \mathbf{a}_j$.

Define $r_{jj} = \|\mathbf{y}_j\|$ and $r_{ij} = \mathbf{q}_i^\top \mathbf{a}_j$.

$$\begin{aligned}
\mathbf{a}_1 &= r_{11}\mathbf{q}_1 \\
\mathbf{a}_2 &= r_{12}\mathbf{q}_1 + r_{22}\mathbf{q}_2 \\
&\vdots \\
\mathbf{a}_j &= r_{1j}\mathbf{q}_1 + r_{2j}\mathbf{q}_2 + \ldots + r_{jj}\mathbf{q}_j
\end{aligned}$$

Define $r_{jj} = \|\mathbf{y}_j\|$ and $r_{ij} = \mathbf{q}_i^\top \mathbf{a}_j$.

$$\begin{aligned}
\mathbf{a}_1 &= r_{11}\mathbf{q}_1 \\
\mathbf{a}_2 &= r_{12}\mathbf{q}_1 + r_{22}\mathbf{q}_2 \\
&\vdots \\
\mathbf{a}_j &= r_{1j}\mathbf{q}_1 + r_{2j}\mathbf{q}_2 + \ldots + r_{jj}\mathbf{q}_j
\end{aligned}$$

$$\underbrace{\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \ldots & \mathbf{a}_n \end{bmatrix}}_{\mathtt{A}_{m\times n}} = \underbrace{\begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \ldots & \mathbf{q}_n \end{bmatrix}}_{\mathtt{Q}_{m\times n}} \underbrace{\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{nn} \end{bmatrix}}_{\mathtt{R}_{n\times n}}$$

Example

Find the reduced QR factorization of $\mathtt{A} = \begin{bmatrix} 1 & -4 \\ 2 & 3 \\ 2 & 2 \end{bmatrix}$.

Find the reduced QR factorization of $\mathtt{A} = \begin{bmatrix} 1 & -4 \\ 2 & 3 \\ 2 & 2 \end{bmatrix}$.

$$\mathbf{y}_1 = \mathbf{a}_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$$

Example

Find the reduced QR factorization of $\mathtt{A} = \begin{bmatrix} 1 & -4 \\ 2 & 3 \\ 2 & 2 \end{bmatrix}$.

$$\mathbf{y}_1 = \mathbf{a}_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \quad \mathbf{q}_1 = \frac{\mathbf{y}_1}{\|\mathbf{y}_1\|} = \frac{\mathbf{y}_1}{3} = \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{bmatrix}$$

Find the reduced QR factorization of $\mathtt{A} = \begin{bmatrix} 1 & -4 \\ 2 & 3 \\ 2 & 2 \end{bmatrix}$.

$$\mathbf{y}_1 = \mathbf{a}_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \quad \mathbf{q}_1 = \frac{\mathbf{y}_1}{\|\mathbf{y}_1\|} = \frac{\mathbf{y}_1}{3} = \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{bmatrix}$$

$$\mathbf{y}_2 = \mathbf{a}_2 - \mathbf{q}_1(\mathbf{q}_1^\top \mathbf{a}_2)$$

Find the reduced QR factorization of $\mathtt{A} = \begin{bmatrix} 1 & -4 \\ 2 & 3 \\ 2 & 2 \end{bmatrix}$.

$$\mathbf{y}_1 = \mathbf{a}_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \quad \mathbf{q}_1 = \frac{\mathbf{y}_1}{\|\mathbf{y}_1\|} = \frac{\mathbf{y}_1}{3} = \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{bmatrix}$$

$$\mathbf{y}_2 = \mathbf{a}_2 - \mathbf{q}_1(\mathbf{q}_1^\top \mathbf{a}_2) = \begin{bmatrix} -4 \\ 3 \\ 2 \end{bmatrix} - \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{bmatrix} \underbrace{\left(\begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} -4 \\ 3 \\ 2 \end{bmatrix} \right)}_{2} = \begin{bmatrix} -\frac{14}{3} \\ \frac{5}{3} \\ \frac{2}{3} \end{bmatrix},$$

Find the reduced QR factorization of $\mathbf{A} = \begin{bmatrix} 1 & -4 \\ 2 & 3 \\ 2 & 2 \end{bmatrix}$.

$$\mathbf{y}_1 = \mathbf{a}_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \quad \mathbf{q}_1 = \frac{\mathbf{y}_1}{\|\mathbf{y}_1\|} = \frac{\mathbf{y}_1}{3} = \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{bmatrix}$$

$$\mathbf{y}_2 = \mathbf{a}_2 - \mathbf{q}_1(\mathbf{q}_1^\top \mathbf{a}_2) = \begin{bmatrix} -4 \\ 3 \\ 2 \end{bmatrix} - \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{bmatrix} \underbrace{\left(\begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} -4 \\ 3 \\ 2 \end{bmatrix} \right)}_{2} = \begin{bmatrix} -\frac{14}{3} \\ \frac{5}{3} \\ \frac{2}{3} \end{bmatrix},$$

$$\mathbf{q}_2 = \frac{\mathbf{y}_2}{\|\mathbf{y}_2\|} = \frac{\mathbf{y}_2}{5} = \begin{bmatrix} -\frac{14}{15} \\ \frac{1}{3} \\ \frac{2}{15} \end{bmatrix}$$

Find the reduced QR factorization of $\mathtt{A} = \begin{bmatrix} 1 & -4 \\ 2 & 3 \\ 2 & 2 \end{bmatrix}$.

$$\mathbf{y}_1 = \mathbf{a}_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \quad \mathbf{q}_1 = \frac{\mathbf{y}_1}{\|\mathbf{y}_1\|} = \frac{\mathbf{y}_1}{3} = \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{bmatrix}$$

$$\mathbf{y}_2 = \mathbf{a}_2 - \mathbf{q}_1(\mathbf{q}_1^\top \mathbf{a}_2) = \begin{bmatrix} -4 \\ 3 \\ 2 \end{bmatrix} - \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{bmatrix} \underbrace{\left(\begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} -4 \\ 3 \\ 2 \end{bmatrix} \right)}_{2} = \begin{bmatrix} -\frac{14}{3} \\ \frac{5}{3} \\ \frac{2}{3} \end{bmatrix},$$

$$\mathbf{q}_2 = \frac{\mathbf{y}_2}{\|\mathbf{y}_2\|} = \frac{\mathbf{y}_2}{5} = \begin{bmatrix} -\frac{14}{15} \\ \frac{1}{3} \\ \frac{2}{15} \end{bmatrix} \to \mathtt{A} = \begin{bmatrix} \frac{1}{3} & -\frac{14}{15} \\ \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{2}{15} \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 0 & 5 \end{bmatrix}$$

Input: $\mathtt{A}_{m\times n} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix}$ with linearly independent columns.

for $j \leftarrow 1$ **to** n **do**
  $\mathbf{y} = \mathbf{a}_j$
  for $i \leftarrow 1$ **to** $j-1$ **do**
    $r_{ij} = \mathbf{q}_i^\top \mathbf{a}_j$
    $\mathbf{y} = \mathbf{y} - r_{ij}\mathbf{q}_i$
  end
  $r_{jj} = \|\mathbf{y}\|$
  $\mathbf{q}_j = \frac{\mathbf{y}}{r_{jj}}$
end

Output: $\mathtt{Q}_{m\times n}$ and $\mathtt{R}_{n\times n}$ such that

$$\mathtt{A}_{m\times n} = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \dots & \mathbf{q}_n \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{nn} \end{bmatrix}$$

$$= \mathtt{Q}_{m\times n}\mathtt{R}_{n\times n}$$