Lecture 13: Orthogonal projections and least squares (Section 3.2-3.3)

Thang Huynh, UC San Diego

2/9/2018

Orthogonal projection onto subspaces

Theorem. Let W be a subspace of $\mathbb{R}^n$. Then, each $\boldsymbol{x}$ in $\mathbb{R}^n$ can be uniquely written as

$$\boldsymbol{x} = \underbrace{\hat{\boldsymbol{x}}}_{\text{in } W} + \underbrace{\boldsymbol{x}^{\perp}}_{\text{in } W^{\perp}}.$$

Orthogonal projection onto subspaces

Theorem. Let W be a subspace of $\mathbb{R}^n$. Then, each $\boldsymbol{x}$ in $\mathbb{R}^n$ can be uniquely written as

$$\boldsymbol{x} = \underbrace{\hat{\boldsymbol{x}}}_{\text{in } W} + \underbrace{\boldsymbol{x}^{\perp}}_{\text{in } W^{\perp}}.$$

- $\hat{\boldsymbol{x}}$ is the **orthogonal projection** of $\boldsymbol{x}$ onto W.

Orthogonal projection onto subspaces

Theorem. Let W be a subspace of $\mathbb{R}^n$. Then, each $\boldsymbol{x}$ in $\mathbb{R}^n$ can be uniquely written as

$$\boldsymbol{x} = \underbrace{\hat{\boldsymbol{x}}}_{\text{in } W} + \underbrace{\boldsymbol{x}^{\perp}}_{\text{in } W^{\perp}}.$$

- $\hat{\boldsymbol{x}}$ is the **orthogonal projection** of $\boldsymbol{x}$ onto W.
- $\hat{\boldsymbol{x}}$ is the point in W closest to $\boldsymbol{x}$.

Orthogonal projection onto subspaces

Theorem. Let W be a subspace of $\mathbb{R}^n$. Then, each $\boldsymbol{x}$ in $\mathbb{R}^n$ can be uniquely written as

$$\boldsymbol{x} = \underbrace{\hat{\boldsymbol{x}}}_{\text{in } W} + \underbrace{\boldsymbol{x}^\perp}_{\text{in } W^\perp}.$$

- $\hat{\boldsymbol{x}}$ is the **orthogonal projection** of $\boldsymbol{x}$ onto W.
- $\hat{\boldsymbol{x}}$ is the point in W closest to $\boldsymbol{x}$.
- If $\boldsymbol{v}_1, \dots, \boldsymbol{v}_m$ is an orthogonal basis of W, then

$$\hat{\boldsymbol{x}} = \left(\frac{\boldsymbol{x} \cdot \boldsymbol{v}_1}{\boldsymbol{v}_1 \cdot \boldsymbol{v}_1}\right)\boldsymbol{v}_1 + \dots + \left(\frac{\boldsymbol{x} \cdot \boldsymbol{v}_m}{\boldsymbol{v}_m \cdot \boldsymbol{v}_m}\right)\boldsymbol{v}_m.$$

Orthogonal projection onto subspaces

Theorem. Let W be a subspace of $\mathbb{R}^n$. Then, each $\boldsymbol{x}$ in $\mathbb{R}^n$ can be uniquely written as

$$\boldsymbol{x} = \underbrace{\hat{\boldsymbol{x}}}_{\text{in } W} + \underbrace{\boldsymbol{x}^{\perp}}_{\text{in } W^{\perp}}.$$

- $\hat{\boldsymbol{x}}$ is the **orthogonal projection** of $\boldsymbol{x}$ onto W.
- $\hat{\boldsymbol{x}}$ is the point in W closest to $\boldsymbol{x}$.
- If $\boldsymbol{v}_1, \dots, \boldsymbol{v}_m$ is an orthogonal basis of W, then

$$\hat{\boldsymbol{x}} = \left(\frac{\boldsymbol{x} \cdot \boldsymbol{v}_1}{\boldsymbol{v}_1 \cdot \boldsymbol{v}_1}\right)\boldsymbol{v}_1 + \dots + \left(\frac{\boldsymbol{x} \cdot \boldsymbol{v}_m}{\boldsymbol{v}_m \cdot \boldsymbol{v}_m}\right)\boldsymbol{v}_m.$$

- Once $\hat{\boldsymbol{x}}$ is determined, $\boldsymbol{x}^{\perp} = \boldsymbol{x} - \hat{\boldsymbol{x}}$.

Orthogonal projection onto subspaces

- Example. Let $W = \text{span}\left\{ \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$, and $\boldsymbol{x} = \begin{bmatrix} 0 \\ 3 \\ 10 \end{bmatrix}$.
  - Find the orthogonal projection of $\boldsymbol{x}$ onto W.
  - Write $\boldsymbol{x}$ as a vector in W plus a vector orthogonal to W.

Orthogonal projection onto subspaces

- Definition. Let $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m$ be an orthogonal basis of W, a subspace of $\mathbb{R}^n$. The projection map $\pi_W : \mathbb{R}^n \to \mathbb{R}^n$, given by

$$\pi_W(\boldsymbol{x}) = \left(\frac{\boldsymbol{x} \cdot \boldsymbol{v}_1}{\boldsymbol{v}_1 \cdot \boldsymbol{v}_1}\right)\boldsymbol{v}_1 + \ldots + \left(\frac{\boldsymbol{x} \cdot \boldsymbol{v}_m}{\boldsymbol{v}_m \cdot \boldsymbol{v}_m}\right)\boldsymbol{v}_m$$

is linear (why?). The matrix P representing π_W with respect to the standard basis is the corresponding **projection matrix.**

Orthogonal projection onto subspaces

- Definition. Let $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m$ be an orthogonal basis of W, a subspace of $\mathbb{R}^n$. The projection map $\pi_W : \mathbb{R}^n \to \mathbb{R}^n$, given by

$$\pi_W(\boldsymbol{x}) = \left(\frac{\boldsymbol{x} \cdot \boldsymbol{v}_1}{\boldsymbol{v}_1 \cdot \boldsymbol{v}_1}\right)\boldsymbol{v}_1 + \ldots + \left(\frac{\boldsymbol{x} \cdot \boldsymbol{v}_m}{\boldsymbol{v}_m \cdot \boldsymbol{v}_m}\right)\boldsymbol{v}_m$$

is linear (why?). The matrix P representing π_W with respect to the standard basis is the corresponding **projection matrix.**

- Example. Find the projection matrix P which corresponds to orthogonal projection onto $W = \text{span}\left\{\begin{bmatrix}3\\0\\1\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}\right\}$ in $\mathbb{R}^3$. Then find the orthogonal projection of $\boldsymbol{x} = \begin{bmatrix}0\\3\\10\end{bmatrix}$ onto W.

Least squares

In practice,

$$A\boldsymbol{x} \approx \boldsymbol{b}.$$

Least squares

In practice,

$$A\boldsymbol{x} \approx \boldsymbol{b}.$$

- Definition. $\hat{\boldsymbol{x}}$ is a **least squares solution** of the system $A\boldsymbol{x} = \boldsymbol{b}$ if $\hat{\boldsymbol{x}}$ is such that $A\hat{\boldsymbol{x}} - \boldsymbol{b}$ is as small as possible.

Least squares

In practice,

$$A\boldsymbol{x} \approx \boldsymbol{b}.$$

- Definition. $\hat{\boldsymbol{x}}$ is a **least squares solution** of the system $A\boldsymbol{x} = \boldsymbol{b}$ if $\hat{\boldsymbol{x}}$ is such that $A\hat{\boldsymbol{x}} - \boldsymbol{b}$ is as small as possible.
- If $A\boldsymbol{x} = \boldsymbol{b}$ is consistent, i.e. $\boldsymbol{b}$ is in $C(A)$, then a least squares solution $\hat{\boldsymbol{x}}$ is just an ordinary solution.

Least squares

In practice,

$$A\boldsymbol{x} \approx \boldsymbol{b}.$$

- Definition. $\hat{\boldsymbol{x}}$ is a **least squares solution** of the system $A\boldsymbol{x} = \boldsymbol{b}$ if $\hat{\boldsymbol{x}}$ is such that $A\hat{\boldsymbol{x}} - \boldsymbol{b}$ is as small as possible.
- If $A\boldsymbol{x} = \boldsymbol{b}$ is consistent, i.e. $\boldsymbol{b}$ is in $C(A)$, then a least squares solution $\hat{\boldsymbol{x}}$ is just an ordinary solution.
- Interesting case: $A\boldsymbol{x} = \boldsymbol{b}$ is inconsistent, i.e. $\boldsymbol{b}$ is NOT in $C(A)$.

Least squares

In practice,

$$A\boldsymbol{x} \approx \boldsymbol{b}.$$

- Definition. $\hat{\boldsymbol{x}}$ is a **least squares solution** of the system $A\boldsymbol{x} = \boldsymbol{b}$ if $\hat{\boldsymbol{x}}$ is such that $A\hat{\boldsymbol{x}} - \boldsymbol{b}$ is as small as possible.
- If $A\boldsymbol{x} = \boldsymbol{b}$ is consistent, i.e. $\boldsymbol{b}$ is in $C(A)$, then a least squares solution $\hat{\boldsymbol{x}}$ is just an ordinary solution.
- Interesting case: $A\boldsymbol{x} = \boldsymbol{b}$ is inconsistent, i.e. $\boldsymbol{b}$ is NOT in $C(A)$. What should we do to find $\hat{\boldsymbol{x}}$?

Least squares

In practice,

$$A\boldsymbol{x} \approx \boldsymbol{b}.$$

▶ Definition. $\hat{\boldsymbol{x}}$ is a **least squares solution** of the system $A\boldsymbol{x} = \boldsymbol{b}$ if $\hat{\boldsymbol{x}}$ is such that $A\hat{\boldsymbol{x}} - \boldsymbol{b}$ is as small as possible.

▶ If $A\boldsymbol{x} = \boldsymbol{b}$ is consistent, i.e. $\boldsymbol{b}$ is in $C(A)$, then a least squares solution $\hat{\boldsymbol{x}}$ is just an ordinary solution.

▶ Interesting case: $A\boldsymbol{x} = \boldsymbol{b}$ is inconsistent, i.e. $\boldsymbol{b}$ is NOT in $C(A)$. What should we do to find $\hat{\boldsymbol{x}}$?

- replace $\boldsymbol{b}$ with its projection $\hat{\boldsymbol{b}}$ onto $C(A)$
- and solve $A\boldsymbol{x} = \hat{\boldsymbol{b}}$.

The normal equations

Theorem $\hat{\boldsymbol{x}}$ is a least squares solution of $A\boldsymbol{x} = \boldsymbol{b}$ if and only if $A^T A\hat{\boldsymbol{x}} = A^T\boldsymbol{b}$ (the normal equation).

Proof.

The normal equations

- Example. Find the least squares solution to $A\boldsymbol{x} = \boldsymbol{b}$, where

$$A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \\ 0 & 0 \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}.$$

The normal equations

- Example. Find the least squares solution to $A\boldsymbol{x} = \boldsymbol{b}$, where

$$A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \\ 0 & 0 \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}.$$

- Solution.

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

and

$$A^T \boldsymbol{b} = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}.$$

The normal equations

The normal equation $A^T A\boldsymbol{x} = A^T\boldsymbol{b}$ is

$$\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}.$$

Solving it, we obtain $\hat{\boldsymbol{x}} = \begin{bmatrix} 1/2 \\ 3/2 \end{bmatrix}$.

The normal equations

- Example. Find the least squares solution to $A\boldsymbol{x} = \boldsymbol{b}$, where

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

What is the projection of $\boldsymbol{b}$ onto $C(A)$?

The normal equations

- Example. Find the least squares solution to $A\boldsymbol{x} = \boldsymbol{b}$, where

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

What is the projection of $\boldsymbol{b}$ onto $C(A)$?

- Solution.

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

and

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

The normal equations

The normal equation $A^T A\boldsymbol{x} = A^T\boldsymbol{b}$ is

$$\begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix}\boldsymbol{x} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}.$$

Solving it, we obtain $\hat{\boldsymbol{x}} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

The normal equations

The normal equation $A^T A\boldsymbol{x} = A^T\boldsymbol{b}$ is

$$\begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}.$$

Solving it, we obtain $\hat{\boldsymbol{x}} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

The projection of $\boldsymbol{b}$ onto $C(A)$ is $A\hat{\boldsymbol{x}} = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ 3 \end{bmatrix}$.

Why is $A\hat{x}$ the projection of b onto $C(A)$?

The projection $\hat{\boldsymbol{b}}$ of $\boldsymbol{b}$ onto $C(A)$ is

$$\hat{\boldsymbol{b}} = A\hat{\boldsymbol{x}}, \quad \text{with } \hat{\boldsymbol{x}} \text{ such that } A^T A\hat{\boldsymbol{x}} = A^T\boldsymbol{b}.$$

Why is $A\hat{x}$ the projection of b onto $C(A)$?

The projection $\hat{\boldsymbol{b}}$ of $\boldsymbol{b}$ onto $C(A)$ is

$$\hat{\boldsymbol{b}} = A\hat{\boldsymbol{x}}, \quad \text{with } \hat{\boldsymbol{x}} \text{ such that } A^T A\hat{\boldsymbol{x}} = A^T\boldsymbol{b}.$$

If A has full column rank (columns of A linearly independent), then

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

The projection matrix for projecting onto $C(A)$ is

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

Application: least squares lines

- Example. Find β_1, β_2 such that the line $y = \beta_1 + \beta_2 x$ best fits the data points $(2, 1), (5, 2), (7, 3), (8, 3)$.

Application: least squares lines

- Example. Find β_1, β_2 such that the line $y = \beta_1 + \beta_2 x$ best fits the data points $(2, 1), (5, 2), (7, 3), (8, 3)$.
- Solution. The equations $y_i = \beta_1 + \beta_2 x_i$ in matrix form

$$\underbrace{\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ 1 & x_4 \end{bmatrix}}_{\text{design matrix } X} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}}_{\text{observation vector } y} .$$

Application: least squares lines

We need to find a least squares solution to

$$\begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix}.$$

Then

$$X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} = \begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix},$$

$$X^T \boldsymbol{y} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix} = \begin{bmatrix} 9 \\ 57 \end{bmatrix} \Longrightarrow \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 2/7 \\ 5/14 \end{bmatrix}.$$