Lecture 24: Sections 6.3, 6.5

Orthogonal projection

Let H be a subspace of $\mathbb{R}^n$. Then for any $\mathbf{x} \in \mathbb{R}^n$, there is unique $\hat{\mathbf{x}} \in H$ such that

$$\mathbf{x} = \hat{\mathbf{x}} + \mathbf{z} \qquad \text{with } \mathbf{z} \perp H.$$

If $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an orthogonal basis of H, then $\mathbf{z} = \mathbf{x} - \hat{\mathbf{x}} \perp H$ with

$$\hat{\mathbf{x}} = \frac{(\mathbf{x} \cdot \mathbf{u}_1)}{(\mathbf{u}_1 \cdot \mathbf{u}_1)}\mathbf{u}_1 + \ldots + \frac{(\mathbf{x} \cdot \mathbf{u}_p)}{(\mathbf{u}_p \cdot \mathbf{u}_p)}\mathbf{u}_p.$$

$\hat{\mathbf{x}}$ is called the orthogonal projection of $\mathbf{x}$ onto H, and denoted by $\text{Proj}_H \mathbf{x} = \hat{\mathbf{x}}$.

- $\mathbf{x} \in H \Rightarrow \text{Proj}_H \mathbf{x} = \mathbf{x}$
- $\|\mathbf{x} - \hat{\mathbf{x}}\| < \|\mathbf{x} - \mathbf{v}\|$ for any $\mathbf{v} \in H$ with $\mathbf{v} \neq \hat{\mathbf{x}} = \text{Proj}_H \mathbf{x}$
- If $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an orthonormal basis of H, then

$$\text{Proj}_H \mathbf{x} = (\mathbf{x} \cdot \mathbf{u}_1)\mathbf{u}_1 + \ldots + (\mathbf{x} \cdot \mathbf{u}_p)\mathbf{u}_p = UU^T\mathbf{x}$$

with the orthogonal matrix $U = [\mathbf{u}_1 \ldots \mathbf{u}_p]$

Least-squares problems

$k \times n$ matrix A

- $A\mathbf{x} = \mathbf{b}$ is inconsistent $\Leftrightarrow$ $\mathbf{b} \notin \operatorname{Col} A$.
- We have $\hat{\mathbf{b}} = \operatorname{Proj}_{\operatorname{Col} A} \mathbf{b} \in \operatorname{Col} A$.

Any solution of $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$ is called a least-squares solution of $A\mathbf{x} = \mathbf{b}$.

- $A\hat{\mathbf{x}} = \hat{\mathbf{b}} \quad \Leftrightarrow \quad \|\mathbf{b} - A\hat{\mathbf{x}}\| \leq \|\mathbf{b} - A\mathbf{x}\|$ for any $\mathbf{x} \in \mathbb{R}^n$ (the least-squares problem)
- $A\hat{\mathbf{x}} = \hat{\mathbf{b}} \quad \Leftrightarrow \quad A^T A\hat{\mathbf{x}} = A^T\mathbf{b}$ (the normal equation)
- The columns of A are linearly independent $\Leftrightarrow$ $A^T A$ is invertible $\Leftrightarrow$ $A^T A\hat{\mathbf{x}} = A^T\mathbf{b}$ has a unique solution

Experimental data: n points (x_i, y_i) on the plane. Fit the data by a line $y = \beta_0 + \beta_1 x$:

$$X\boldsymbol{\beta} = \mathbf{y} \qquad \text{with} \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}, \quad X = \begin{bmatrix} 1 & x_1 \\ \dots & \dots \\ 1 & x_n \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ \dots \\ y_n \end{bmatrix}$$

The least-squares solution minimizes $\|\mathbf{y} - X\boldsymbol{\beta}\|^2 = \sum_{i=1}^{n} \left[y_i - (\beta_0 + \beta_1 x_i)\right]^2$