CPSC 536F November 5, 2025

Variational Formulas:

$$f(\vec{u}) = \|A\vec{u} - \vec{b}\|^2, \quad A \in \mathbb{R}^{m \times n} \text{ arbitrary}$$

$$f(\vec{u} + \varepsilon\vec{v}) = f(\vec{u}) + \varepsilon \, \mathrm{Term}_1 + \varepsilon^2 \, \mathrm{Term}_2$$

where

$$\mathrm{Term}_1 = 2(A\vec{v}) \cdot (A\vec{u} - \vec{b})$$

$$= 2\,\vec{v} \cdot \underbrace{\left(A^T(A\vec{u} - \vec{b})\right)}_{\text{Setting to 0} \Rightarrow \text{Normal equation}}$$

$$g(\vec{u}) = (A\vec{u}) \cdot \vec{u} \qquad A \in \mathbb{R}^{n \times n} \text{ symmetric}$$

$$g(\vec{u} + \varepsilon\vec{v}) = g(\vec{u}) + \varepsilon \, \mathrm{Term}_1 + \varepsilon^2 \, \mathrm{Term}_2$$

$$\mathrm{Term}_1 = 2\,\vec{v} \cdot (A\vec{u})$$

Algebra:

$$\frac{\alpha_0 + \alpha_1 \varepsilon + O(\varepsilon^2)}{\beta_0 + \beta_1 \varepsilon + O(\varepsilon^2)} = \frac{\alpha_0}{\beta_0} + \varepsilon \, \mathrm{Term}_1 + O(\varepsilon^2)$$

where

$$\mathrm{Term}_1 = 0 \iff \alpha_0 \beta_1 = \alpha_1 \beta_0$$

$$\left[\mathrm{Term}_1 = \frac{\alpha_0}{\beta_0} \left(\frac{\alpha_1}{\alpha_0} - \frac{\beta_1}{\beta_0} \right) \right]$$

Rem: If $\beta_1 = 0$, Then $\mathrm{Term}_1 = 0 \iff \alpha_1 = 0$

Exercise: Say $\beta_1 = 0$, what are Term_1, Term_2 in

$$\frac{\alpha_0 + \alpha_1 \varepsilon + \alpha_2 \varepsilon^2 + O(\varepsilon^3)}{\beta_0 + \qquad \beta_2 \varepsilon^2 + O(\varepsilon^3)} \quad ??$$

Good news: All Sept homeworks were done very well, the writing was crystal clear.

Specific comments Friday or during the reading break

Exercise sections: A, B, C done

D in progress

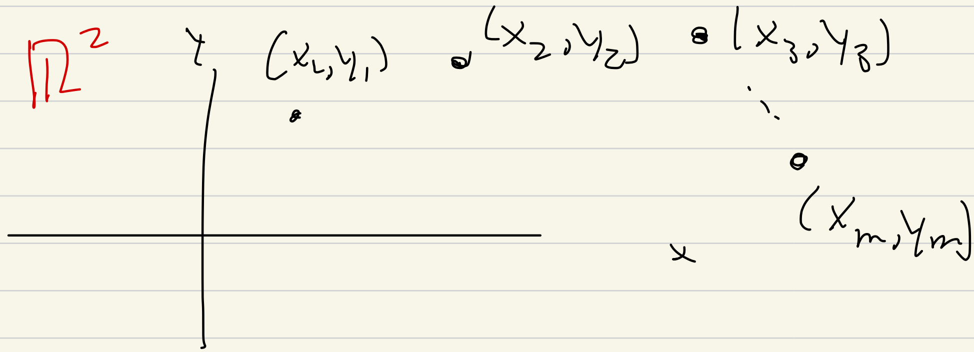
E : apps of spectral decomp.

2 applications

(1) $$R_A(\vec{u}) = \frac{(A\vec{u})\cdot\vec{u}}{\vec{u}\cdot\vec{u}}$$

(2) $$y = u_1 + u_2 x + u_3 x^2$$

find best $u_1, u_2, u_3 \in \mathbb{R}$ that is a good model for data



$\forall i \in [m]$

$$y_i \approx u_1 + u_2 x_i + u_3 x_i^2$$

approx as good as possible

"least squares"

$$E = \mathrm{Error}(u_1, u_2, u_3)$$

$$= \sum_{i=1}^{m} \left(y_i - (u_1 + u_2 x_i + u_3 x_i^2) \right)^2$$

$$= \Bigg\| \underbrace{\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}}_{\vec{y}_0} - u_1 \underbrace{\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}}_{\mathbb{1}} - u_2 \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}}_{\vec{x}_0} - u_3 \underbrace{\begin{bmatrix} x_1^2 \\ x_2^2 \\ \vdots \\ x_m^2 \end{bmatrix}}_{\vec{x}_0^2} \Bigg\|^2$$

$$m = \sum_{i=1}^{m} 1$$

Normal Equations

$$\begin{bmatrix} \sum 1 & \sum x_i & \sum x_i^2 \\ \sum x_i & \sum x_i^2 & \sum x_i^3 \\ \sum x_i & \sum x_i^3 & \sum x_i^4 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} \sum x_i \cdot 1 \\ \sum x_i \cdot y_i \\ \sum x_i \cdot y_i^2 \end{bmatrix}$$

$$\sum = \sum_{i=1}^{m} \text{, i.e.}$$

$$\begin{bmatrix} - & \vec{\mathbb{1}} & - \\ - & \vec{x}_0 & - \\ - & \vec{x}_0^{\,2} & - \end{bmatrix} \begin{bmatrix} | & | & | \\ \vec{\mathbb{1}}^T & \vec{x}_1^{\,T} & \vec{x}_0^{\,2\,T} \\ | & | & | \end{bmatrix}$$

$$\begin{bmatrix} - & \mathbb{1} & - \\ - & \vec{x}_0 & - \\ - & \vec{x}_0^{\,2} & - \end{bmatrix} \begin{bmatrix} | \\ \vec{y}^T \\ | \end{bmatrix} = \begin{bmatrix} \sum x_i \cdot 1 \\ \sum x_i \cdot y_i \\ \sum x_i \cdot y_i^2 \end{bmatrix}$$

Can easily discard, add, then

$$\left(\overset{\text{slightly}}{\text{different}} \right) \vec{u} = \begin{bmatrix} \end{bmatrix}$$

slightly different

3x3 system

Claim: This is minimizing $\| A\vec{u} - \vec{b} \|^2$

$$A = \begin{bmatrix} | & | & | \\ \mathbb{1}^T & \vec{x}_0^T & (\vec{x}^2)^T \\ | & | & | \end{bmatrix}, \quad \vec{b} = \vec{y}_0^{\,T}$$

So

$$\underset{\vec{u} \in \mathbb{R}^n}{\text{minimize}} \; \| A\vec{u} - \vec{b} \|_{L^2}$$

$\Longrightarrow$ $\vec{u}$ is a minimum

then $\forall \vec{v} \in \mathbb{R}^n$

$$\vec{v} \cdot \left(A^T (A\vec{u} - \vec{b}) \right) = 0$$

then we must have

$$A^T \left(A\vec{u} - \vec{b} \right) = 0$$

$$\Rightarrow \quad (A^T A)\vec{u} = A^T \vec{b}$$

$C = 0$ regression

Ridge Regression:

$\rightarrow C$ large

$$f(\vec{u}) = \|A\vec{u} - \vec{b}\|^2 + C\|\vec{u}\|^2$$

$$\min_{\vec{u} \in \mathbb{R}^d} f(\vec{u})$$

$$\vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \\ u_d \end{bmatrix}$$

$d > 3$

$d \geq 10$

$$\tilde{\hat{A}}\vec{u}-\tilde{\hat{b}}$$

$$f(\vec{u}) = \left\| A\vec{u}-\vec{b} \right\|^2 + C \left\| I\cdot\vec{u}-\vec{0} \right\|^2$$

$$\hat{A} = I \qquad \tilde{\hat{b}} = 0$$

we already know

$$f(\vec{u}+\varepsilon\vec{V}) = f(\vec{u}) + \varepsilon \, Term_1 + O(\varepsilon^2)$$

$$Term_1 = 2\vec{V}\cdot\left(A^T(A\vec{u}-\vec{b})\right) + 2\vec{V}\cdot\left(I^T(I\vec{u}-0)\right)$$

$$\min/_{max} \Rightarrow \left.\begin{matrix} A^T A\vec{u} - A^T b \\ + C(\vec{u}-\vec{0}) \end{matrix}\right\} = \vec{0}$$