EE16A - Lecture 22 Notes

Name: Felix Su   SID: 25794773

Spring 2016   GSI: Ena Hariyoshi

Least Squares Geometric Argument

> 1. Best estimate (1D): $\hat{b} = \frac{\vec{a}^T\vec{b}}{\vec{a}^T\vec{a}}\vec{a}$ where $\frac{\vec{a}^T\vec{b}}{\vec{a}^T\vec{a}}$= best scalar x
> 2. Best estimate (Multi-D): $x = (A^T A)^{-1} A^T b$
> 3. Error Vector: $\varepsilon = \vec{b} - A\vec{x}$ because $\vec{b} \notin$ span$\{\vec{a_1}, \ldots, \vec{a_n}\}$
> 4. Goal: Minimize $\|\varepsilon\|^2$ over all possible $\vec{x}$
> 5. Best $\vec{x}$ ensures that $\varepsilon \perp \sum_{i=1}^{n} \alpha_i \vec{a}_i \forall \alpha_i \in \mathbb{R} \implies \varepsilon \perp a_n$ where a_n = each column vector of A
> 6. $(A^T A)^{-1}$ exists iff A has full col. rank (if $A^T A$ is invertible, can't solve), $n >> m \implies$ this will happen.

- Dim(A)=$n \times m$, Dim(x)=$m \times 1$, Dim(b)=$n \times 1$
  - Generally $n > m$
- **Claim:** $A^T A$ is invertible iff A has linearly independent columns.
- **Proof:** Show Null$(A^T A)$ = Null(A)
  - $A^T A$ guaranteed to be square $m \times m$
  - Forward: $\vec{q} \in$ Null$(A), \exists \vec{q} \neq 0 \implies \vec{q} \in$ Null$(A^T A)$
    * Because $A\vec{q} = 0 \implies A^T A\vec{q} = A^T 0 = 0$
  - Backward (Converse): $\vec{r} \in$ Null$(A^T A), \exists \vec{r} \neq 0 \implies \vec{r} \in$ Null(A)
    * Because $A^T A\vec{r} = 0 \implies \vec{r}^T A^T A\vec{r} = \langle A\vec{r}, A\vec{r}\rangle = \|A\vec{r}\|^2 = \vec{r}^T 0 = 0$ (**Transpose of a dot product reverses the order of the operands in the product**)
    * $\|A\vec{r}\|^2 = 0 \implies A\vec{r} = 0 \implies \vec{r} \in$ Null(A)
  - Proves Null$(A^T A)$ and Null(A) have the same null space.
- Only vector in null space of Null$(A^T A)$ is $\vec{0} \implies$ columns of A are linearly independent.

Inverse Formula for 2×2 Matrix:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \tag{1}$$

Line of Best Fit

- Example: $t_1 = 0, t_2 = 1, t_2 = 2$ and $y(t) = -1, y(t) = -3, y(t) = 1$, Find $y(t) = x_1 t + x_2$
- Set up $A\vec{x} = y(\vec{t_n})$
- Solve for $A^T A$

Line of Best Fit Example

$$\begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} y(t_1) = -1 \\ y(t_2) = -3 \\ y(t_3) = 1 \end{bmatrix} \tag{2}$$

Find $A^T A$

$$\begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 3 \\ 3 & 3 \end{bmatrix} \tag{3}$$

Find $(A^T A)^{-1}$

$$\frac{1}{6} \begin{bmatrix} 3 & -3 \\ -3 & 5 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{5}{6} \end{bmatrix} \tag{4}$$

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

$$\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{5}{6} \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ -3 \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} & 0 & \frac{1}{2} \\ \frac{5}{6} & \frac{2}{3} & \frac{1}{6} \end{bmatrix} \begin{bmatrix} -1 \\ -3 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \end{bmatrix} \tag{5}$$

Plug in to solve for $y(t)$

$$y(t) = x_1 t + x_2 = t - 2 \tag{6}$$

Find $\varepsilon = \vec{b} - A\vec{x}$

$$\varepsilon = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \tag{7}$$

Determine $\|\varepsilon\|^2$

$$\|\varepsilon\|^2 = 1^2 + (-2)^2 + 1^2 = 6 \tag{8}$$

Least Squares Calculus Argument

> 1. Smallest $\|\varepsilon\|^2$ is the sum of the squares of the distance between each actual measurement (point) and the line of best fit.

Least Squares Calculus:

$$\|\varepsilon\|^2 = \sum_{i=1}^{n(=3)} \varepsilon^2 = \sum_{i=1}^{n(=3)} (x_1 t_i + x_2 - y(t_i))^2 = (x_2+1)^2 + (x_1+x_2+3)^2 + (2x_1+x_2-1)^2 \tag{9}$$

Find minimum w.r.t x_1:

$$\frac{\partial \|\varepsilon\|}{\partial x_1} = 0 + 2(x_1+x_2+3) + 4(2x_1+x_2-1) = 0 \implies 10x_1 + 6x_2 = -2 \tag{10}$$

Find minimum w.r.t x_2:

$$\frac{\partial \|\varepsilon\|}{\partial x_2} = 2(x_2+1) + 2(x_1+x_2+3) + 2(2x_1+x_2-1) = 0 \implies x_1 + x_2 = -1 \tag{11}$$

Solve for x_1, x_2:

$$x_1 = -x_2 - 1, 10(-x_2-1) + 6x_2 = -2, x_2 = -2, x_1 = 1 \tag{12}$$