

PC: Feature Transform

Song Liu

Polynomial Feature Transform

- Given i, b and d , The i -th dimension of polynomial feature transform $\mathbf{f}(\mathbf{x})$ has the form $f^{(i)}(\mathbf{x}) = (x^{(m)})^n$. Express m and n using i, b and d .
- $m = \left\lfloor \frac{i-1}{b+1} \right\rfloor + 1, n = (i-1) \% (b+1)$
 - $\left\lfloor \frac{a}{b} \right\rfloor, a \% b$ are quotient and remainder.

Computational Complexity

- Computational complexity of the LS solution of $\hat{\beta}$ using $f(\mathbf{x}) \in R^m$?

$$\bullet \hat{\beta} = \underbrace{(f(X)^{\top} f(X))^{-1}}_A \underbrace{f(X)^{\top} \mathbf{y}}_B$$

- $A = f(X)^{\top} f(X): O(m \times n \times m)$

- $(A)^{-1}: O(m^3)$

- $B = f(X)^{\top} \mathbf{y}: O(m \times n \times 1)$

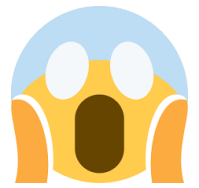
- $(A)^{-1} B: O(m \times m \times 1)$

- **Fixing n , comp. complexity is $O(m^3)$.**

Expected Error

- Rewrite $\mathbb{E}_{\epsilon}[(y - \hat{y}_i)^2 | \mathbf{x}_i]$ using integral.
 - Conditional Expectation:
 - $\mathbb{E}[g(y) | x_i] = \int p(y | x) g(y) dy.$
- $$\begin{aligned}\mathbb{E}_{\epsilon}[(y - \hat{y})^2 | \mathbf{x}_i] &= \int p(y | \mathbf{x}_i) (y - \hat{y})^2 dy \\ &= \int p(y | \mathbf{x}_i) (y - \langle \hat{\boldsymbol{\beta}}, \mathbf{f}(\mathbf{x}_i) \rangle)^2 dy \\ &= \int p(y | \mathbf{x}_i) (y - \langle (\mathbf{F}^{\top} \mathbf{F})^{-1} \mathbf{F}^{\top} \mathbf{y}, \mathbf{f}(\mathbf{x}_i) \rangle)^2 dy\end{aligned}$$
- Note the difference between y and \mathbf{y} .

Expanding Variance Term



- Show $\text{var}[\hat{y}|\mathbf{x}_i]$ is $< h(\mathbf{x}_i), h(\mathbf{x}_i) > \cdot \sigma^2$
 - $h(\mathbf{x}_i) := \mathbf{f}(\mathbf{x}_i)(\mathbf{f}(X)^\top \mathbf{f}(X))^{-1} \mathbf{f}(X)^\top$
- **Hint:**
 - $y = \mathbf{f}(\mathbf{x}_i) + \epsilon$
 - Shorten $\mathbf{f}(\mathbf{x})$ as \mathbf{f} , $\mathbf{f}(X)$ as \mathbf{F} , $\epsilon = [\epsilon_1, \epsilon_2 \dots \epsilon_n]^\top$.
 - $\mathbb{E}_{\mathbf{x}}[\mathbf{x} \mathbf{A} \mathbf{x}^\top] = \text{tr}[\mathbf{A} \Sigma]$, Σ is the covar. of \mathbf{x} .
 - Live demonstration

Induced Kernel

- Write down induced $f(\mathbf{x})$ by polynomial kernels $b=2$.

- $k(\mathbf{x}_i, \mathbf{x}_j) = (\langle \mathbf{x}_i, \mathbf{x}_j \rangle + 1)^b$

- **Hint:**

- Express k using inner product of two vec.

$$k(\mathbf{x}_i, \mathbf{x}_j) = [f^{(1)}(\mathbf{x}) \dots f^{(m)}(\mathbf{x})] \cdot \begin{bmatrix} f^{(1)}(\mathbf{x}) \\ \vdots \\ f^{(m)}(\mathbf{x}) \end{bmatrix}$$

- $f(\mathbf{x}) =$
 $\left[1, x^{(1)} \dots x^{(d)}, (x^{(1)})^2 \dots (x^{(d)})^2, \forall_{u \neq v} x^{(u)} x^{(v)} \right]$