Linear regression (blended lecture)

DS-GA 1013 / MATH-GA 2824 Mathematical Tools for Data Science

Carlos Fernandez-Granda

Review questions on MSE estimation

Linear minimum MSE estimator

Singular-value decomposition

Questions

- What is the minimum MSE estimate of $\tilde{y}$ given $\tilde{x}$?
- Is it easy to approximate?
- Do you know any methods for nonlinear regression?

Questions

- What is the minimum MSE estimate of $\tilde{y}$ given $\tilde{x}$? $\mathrm{E}(\tilde{y} \mid \tilde{x})$
- Is it easy to approximate?
- Do you know any methods for nonlinear regression?

Questions

- What is the minimum MSE estimate of $\tilde{y}$ given $\tilde{x}$? $\mathrm{E}(\tilde{y} \mid \tilde{x})$

- Is it easy to approximate?
  No, unless possible values of $\tilde{x}$ are very restricted

- Do you know any methods for nonlinear regression?

Questions

- What is the minimum MSE estimate of $\tilde{y}$ given $\tilde{x}$? $\mathrm{E}(\tilde{y} \mid \tilde{x})$
- Is it easy to approximate?
  No, unless possible values of $\tilde{x}$ are very restricted
- Do you know any methods for nonlinear regression?
  Nearest neighbors, kernel regression, neural networks

Review questions on MSE estimation

Linear minimum MSE estimator

Singular-value decomposition

Helicopter

Goal: Recording voice of pilot in helicopter

Voice modeled by zero-mean random variable $\tilde{y}$ with variance 1

Microphone inside helmet:

$$\tilde{x}[1] = \tilde{y} + \alpha \tilde{z}$$

Microphone outside helmet:

$$\tilde{x}[2] = \alpha \tilde{y} + \tilde{z}$$

Noise in helicopter modeled by zero-mean random variable $\tilde{z}$ with variance 100 uncorrelated from $\tilde{y}$

Linear minimum MSE estimation

Linear minimum MSE estimator of $\tilde{y}$ given $\tilde{x}$

$$\beta^* = \Sigma_{\tilde{x}}^{-1} \Sigma_{\tilde{x}\tilde{y}}$$

Corresponding MSE

$$\mathrm{E}\left[(\tilde{y} - \tilde{x}^T \Sigma_{\tilde{x}}^{-1} \Sigma_{\tilde{x}\tilde{y}})^2 \right] = \mathrm{Var}(\tilde{y}) - \Sigma_{\tilde{x}\tilde{y}}^T \Sigma_{\tilde{x}}^{-1} \Sigma_{\tilde{x}\tilde{y}}$$

Linear minimum MSE estimate of $\tilde{y}$ given $\tilde{x}[1]$

$$\tilde{x}[1] = \tilde{y} + \alpha\tilde{z}$$

$$\hat{y}(\tilde{x}[1]) =$$

Linear minimum MSE estimate of $\tilde{y}$ given $\tilde{x}[1]$

$$\tilde{x}[1] = \tilde{y} + \alpha \tilde{z}$$

$$\begin{aligned} \hat{y}(\tilde{x}[1]) &= \tilde{x}[1] \, \Sigma_{\tilde{x}[1]}^{-1} \Sigma_{\tilde{x}[1]\tilde{y}} \\ &= \frac{\tilde{x}[1]}{1 + 100\alpha^2} \end{aligned}$$

Additive model

Assume additive noise with zero mean $\tilde{y} = \tilde{x}^T \beta_{\text{true}} + \tilde{w}$

$$\text{MSE} = \mathrm{Var}(\tilde{w})$$

In our case $\tilde{y} = \tilde{x}[1] - \alpha \tilde{z}$ so the MSE is

$$\mathrm{Var}(\alpha \tilde{z}) = 100\alpha^2$$

Is this correct?

MSE

$$\tilde{x}[1] = \tilde{y} + \alpha \tilde{z}$$

MSE =

MSE

$$\tilde{x}[1] = \tilde{y} + \alpha \tilde{z}$$

$$\text{MSE} = \mathrm{Var}(\tilde{y}) - \Sigma_{\tilde{x}[1]\tilde{y}}^{T} \Sigma_{\tilde{x}[1]}^{-1} \Sigma_{\tilde{x}[1]\tilde{y}} = \frac{100\alpha^2}{1 + 100\alpha^2}$$

Linear minimum MSE estimate of $\tilde{y}$ given $\tilde{x}$

$$\tilde{x}[1] = \tilde{y} + \alpha \tilde{z} \qquad \tilde{x}[2] = \alpha \tilde{y} + \tilde{z}$$

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

$$\hat{y}(\tilde{x}) =$$

Linear minimum MSE estimate of $\tilde{y}$ given $\tilde{x}$

$$\tilde{x}[1] = \tilde{y} + \alpha \tilde{z} \qquad \tilde{x}[2] = \alpha \tilde{y} + \tilde{z}$$

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

$$\hat{y}(\tilde{x}) = \tilde{x}^T \Sigma_{\tilde{x}}^{-1} \Sigma_{\tilde{x}\tilde{y}} = \frac{\tilde{x}[1] - \alpha \tilde{x}[2]}{1 - \alpha^2} \qquad \text{MSE?}$$

Review questions on MSE estimation

Linear minimum MSE estimator

Singular-value decomposition

Computing SVD from random inputs

We observe $\tilde{y} := A\tilde{x}$, where $\tilde{x}$ is zero mean and $\Sigma_{\tilde{x}} = I$

How can we compute the singular values of A?

Adversarial perturbation

$$\beta_{\text{OLS}} := \arg\min_{\beta} \|y - X^T\beta\|_2^2$$

What $z \in \mathbb{R}^n$ ($\|z\|_2 \leq \gamma$) perturbs β_{OLS} the most when added to y?

$$\beta_{\text{OLS}}^{\text{mod}}(z) := \arg\min_{\beta} \left|\left|y + z - X^T\beta\right|\right|_2^2$$

$$z_{\text{OLS}}^{\text{adv}} := \arg\max_{||z||_2 \leq \gamma} \left|\left|\beta_{\text{OLS}} - \beta_{\text{OLS}}^{\text{mod}}(z)\right|\right|_2^2$$

Adversarial perturbation

$$\beta_{\text{OLS}} := \arg\min_{\beta} \|y - X^T\beta\|_2^2$$

What $z \in \mathbb{R}^n$ ($\|z\|_2 \leq \gamma$) perturbs β_{OLS} the most when added to y?

$$\beta_{\text{OLS}}^{\text{mod}}(z) := \arg\min_{\beta} \left|\left|y + z - X^T\beta\right|\right|_2^2$$

$$z_{\text{OLS}}^{\text{adv}} := \arg\max_{||z||_2 \leq \gamma} \left|\left|\beta_{\text{OLS}} - \beta_{\text{OLS}}^{\text{mod}}(z)\right|\right|_2^2 = \gamma v_p$$