

MMSE Estimation

Estimation given a PDF

- ▶ Suppose x is an $\mathbb{R}^n$ -valued random variable with pdf p^x .
- ▶ One can *predict* or *estimate* the outcome as follows
 - ▶ Given *cost function* $c : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$
 - ▶ pick *estimate* $\hat{x}$ to minimize $\mathbf{E} c(x, \hat{x})$
- ▶ We will look at the cost function

$$c(x, \hat{x}) = \|x - \hat{x}\|^2$$

- ▶ Then the *mean square error* (MSE) is

$$\mathbf{E}(\|x - \hat{x}\|^2) = \int \|x - \hat{x}\|^2 p^x(x) dx$$

Minimizing the MSE

- ▶ let's find the *minimum mean-square error* (MMSE) estimate of x ; we need to solve

$$\min_{\hat{x}} \mathbf{E}(\|x - \hat{x}\|^2)$$

- ▶ we have

$$\begin{aligned}\mathbf{E}(\|x - \hat{x}\|^2) &= \mathbf{E}((x - \hat{x})^T (x - \hat{x})) \\ &= \mathbf{E}(x^T x - 2\hat{x}^T x + \hat{x}^T \hat{x}) \\ &= \mathbf{E}\|x\|^2 - 2\hat{x}^T \mathbf{E} x + \hat{x}^T \hat{x}\end{aligned}$$

- ▶ differentiating with respect to $\hat{x}$ gives the optimal estimate

$$\hat{x}_{\text{mmse}} = \mathbf{E} x$$

The MMSE estimate

- ▶ the minimum mean-square error estimate of x is

$$\hat{x}_{\text{mmse}} = \mathbf{E} x$$

- ▶ its mean square error is

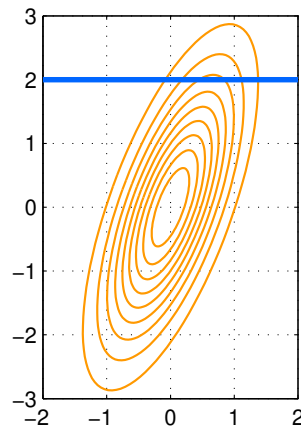
$$\mathbf{E}(\|x - \hat{x}_{\text{mmse}}\|^2) = \text{trace cov}(x) = \sum_{i=1}^n \text{var}(x_i)$$

- ▶ since $\mathbf{E}(\|x - \hat{x}_{\text{mmse}}\|^2) = \mathbf{E}(\|x - \mathbf{E} x\|^2)$

The estimation problem

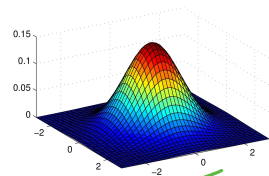
- ▶ Suppose x, y are random variables, with joint pdf $p(x, y)$.
- ▶ We measure $y = y_{\text{meas}}$.
- ▶ We would like to find the MMSE estimate of x given $y = y_{\text{meas}}$.
- ▶ The *estimator* is a function $\phi : \mathbb{R}^m \rightarrow \mathbb{R}^n$.
- ▶ We measure $y = y_{\text{meas}}$, and estimate x by $\hat{x}_{\text{est}} = \phi(y_{\text{meas}})$
- ▶ We would like to find the function ϕ which minimizes the cost function

$$J = \mathbf{E}(\|\phi(y) - x\|^2)$$



Notation

$$p(x, y) \rightarrow \int_y \int_x p(x, y) dx dy = 1$$



$$P(x \leq X \leq x+dx) = f_X(x) dx$$

► We'll use the following notation

► p^y is the *marginal* or *induced* pdf of y

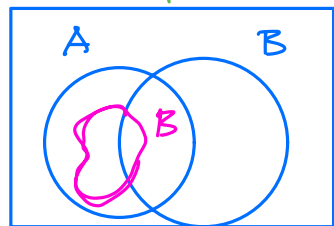
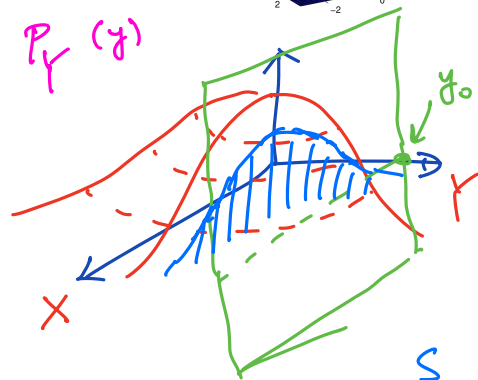
dummy

$$p^y(y) = \int p(x, y) dx$$

P.V.

► $p^{|y}$ is the pdf *conditioned* on y

$$p^{|y}(x, y) = \frac{p(x, y)}{p^y(y)}$$



$$P(A|B) = \frac{P(A \cap B)}{P(B)} = 1$$

$$\int_x p_{|Y=y_0}(x, y_0) dx = 1$$

The MMSE estimator

- ▶ The *mean-square-error conditioned on y* is $e_{\text{cond}}(y)$, given by

$$e_{\text{cond}}(y) = \int \|\phi(y) - x\|^2 p^{\text{ly}}(x, y) dx$$

- ▶ Then the mean square error J is given by

$$J = \mathbf{E}(e_{\text{cond}}(y))$$

- ▶ because

$$\begin{aligned} J &= \int \int \|\phi(y) - x\|^2 p(x, y) dx dy \\ &= \int p^y(y) e_{\text{cond}}(y) dy \end{aligned}$$

The MMSE estimator

- ▶ We can write the MSE conditioned on y as

$$e_{\text{cond}}(y_{\text{meas}}) = \mathbf{E}(\|\phi(y) - x\|^2 \mid y = y_{\text{meas}})$$

- ▶ For each y_{meas} , we can pick a value for $\phi(y_{\text{meas}})$
- ▶ So we have an MMSE prediction problem for each y_{meas}

The MMSE estimator

- recall the mean-variance decomposition

$$\mathbf{E}(\|x\|^2) = \mathbf{E}(\|x - \mathbf{E} x\|^2) + \|\mathbf{E} x\|^2$$

- Apply the MVD to $z = \phi(y) - x$ conditioned on $y = w$ to give

$$\begin{aligned} e_{\text{cond}}(w) &= \mathbf{E}(\|\phi(y) - x\|^2 \mid y = w) \\ &= \mathbf{E}(\|x - h(w)\|^2 \mid y = w) + \|\phi(w) - h(w)\|^2 \end{aligned}$$

where $h(w)$ is the mean of x conditioned on $y = w$

$$h(w) = \mathbf{E}(x \mid y = w)$$

- To minimize $e_{\text{cond}}(w)$ we therefore pick

$$\phi(w) = h(w)$$

The error of the MMSE estimator

- With this choice of estimator

$$\begin{aligned} e_{\text{cond}}(w) &= \mathbf{E}(\|x - h(w)\|^2 \mid y = w) \\ &= \text{trace cov}(x \mid y = w) \end{aligned}$$

The MMSE estimator

- ▶ we have

$$\begin{aligned}\phi_{\text{mmse}}(y_{\text{meas}}) &= \mathbf{E}(x \mid y = y_{\text{meas}}) \\ e_{\text{cond}}(y_{\text{meas}}) &= \mathbf{trace\ cov}(x \mid y = y_{\text{meas}})\end{aligned}$$

- ▶ The estimate only depends on the *conditional pdf* of $x \mid y = y_{\text{meas}}$
- ▶ The means and covariance are those of the *conditional pdf*
- ▶ The above formulae give the MMSE estimate for *any pdf* on x and y

MMSE estimation for Gaussians

$$E(x_1 | x_2 = y) = \Sigma_{12} \Sigma_{22}^{-1} y$$

Suppose $\begin{bmatrix} x \\ y \end{bmatrix} \sim \mathcal{N}(\mu, \Sigma)$ where

$$\text{cov}(x_1 | x_2 = y) = \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}$$

$$\mu = \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix} \quad \Sigma = \begin{bmatrix} \Sigma_x & \Sigma_{xy} \\ \Sigma_{xy}^T & \Sigma_y \end{bmatrix}$$

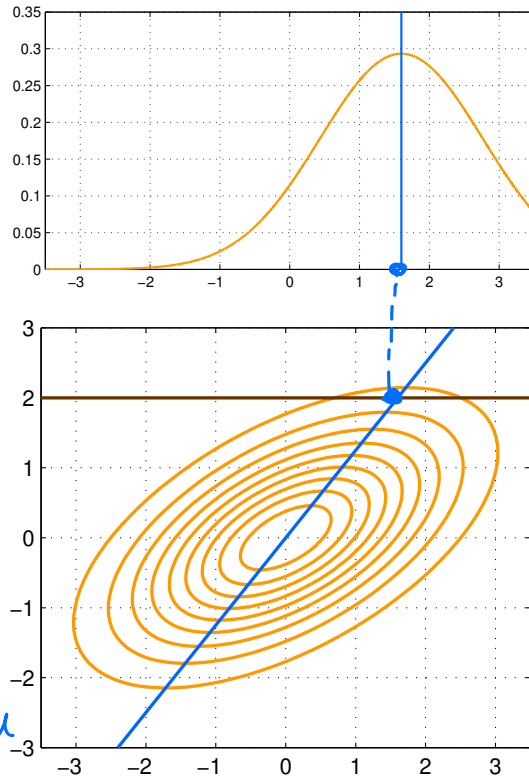
We know that the conditional density of $x | y = y_{\text{meas}}$ is $\mathcal{N}(\mu_1, \Sigma_1)$ where

$$\mu_x + \rho \frac{\sigma_x}{\sigma_y} (y_{\text{meas}} - \mu_y)$$

$$\mu_1 = \mu_x + \Sigma_{xy} \Sigma_y^{-1} (y_{\text{meas}} - \mu_y)$$

$$\Sigma_1 = \Sigma_x - \Sigma_{xy} \Sigma_y^{-1} \Sigma_{xy}^T$$

$$X \rightarrow \tilde{X} = X - \mu$$



MMSE estimation for Gaussians

- ▶ The MMSE estimator when x, y are jointly Gaussian is

$$\hat{x} = \phi_{\text{mmse}}(y_{\text{meas}}) = \mu_x + \Sigma_{xy} \Sigma_y^{-1} (y_{\text{meas}} - \mu_y)$$

- ▶ with error

$$\mathbf{E}(\|x - \hat{x}_{\text{mmse}}\|^2 \mid y = y_{\text{meas}}) = \mathbf{trace}(\Sigma_x - \Sigma_{xy} \Sigma_y^{-1} \Sigma_{xy}^T)$$

- ▶ The conditional mean square error is independent of y ; a special property of Gaussians
- ▶ The estimate $\hat{x}_{\text{mmse}}$ is an *affine* function of y_{meas}