ECE 645

J. V. Krogmeier

Session 2.A – January 22, 2021

Statistical Signal Proc. in a Nutshell

We make observations of some random phenom.

$$Y = y$$

Try to make a model to explain

\# of possible models

finite → Detection Theory

infinite → Estimation Theory

Char. of Observations : Y is a r.v. y is a value it takes.

- $Y = y \in \mathbb{R}, \mathbb{R}^n, \mathbb{C}$ ...
- $Y = y \in \mathbb{D}$ finite or Countable set.

Need notion of Events...

FORM C
APPROVED FOR USE IN
PURDUE UNIVERSITY

Space of Observations $(\Gamma, \mathcal{G})$

Γ: Euclidean or Discrete

$\mathcal{G}$: events assoc. with Γ.
$\mathcal{B}$ Borel sets
2^{Γ} power set.

Parameter Space Λ

$\Lambda \rightarrow$ indexing states of nature
Euclidean or Discrete

Family of Prob. Distributions $\{P_\theta : \theta \in \Lambda\}$

Specifying pdfs $f_\theta(y)$
pmf $p_\theta(y)$

Borel Sets

$X : \Omega \to \mathbb{R}$ a real-valued r.v.

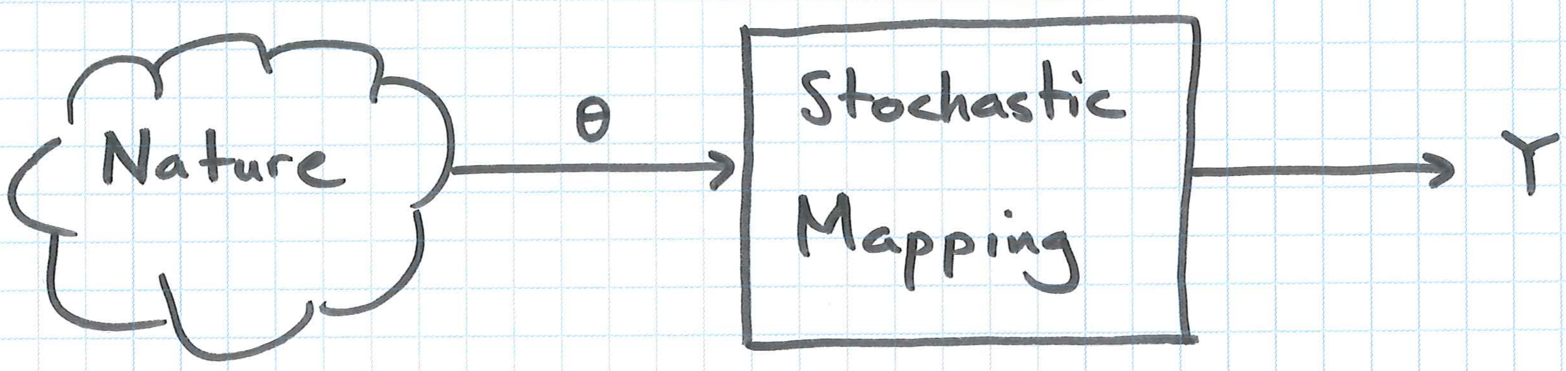
$\mathcal{B}$ is smallest σ-field containing all sets of form $(-\infty, x]$ $x \in \mathbb{R}$

$\mathcal{B}$ contains $\longrightarrow$ all intervals
all open sets
all closed sets

$$P_X\big((-\infty, x]\big) = F_X(x)$$

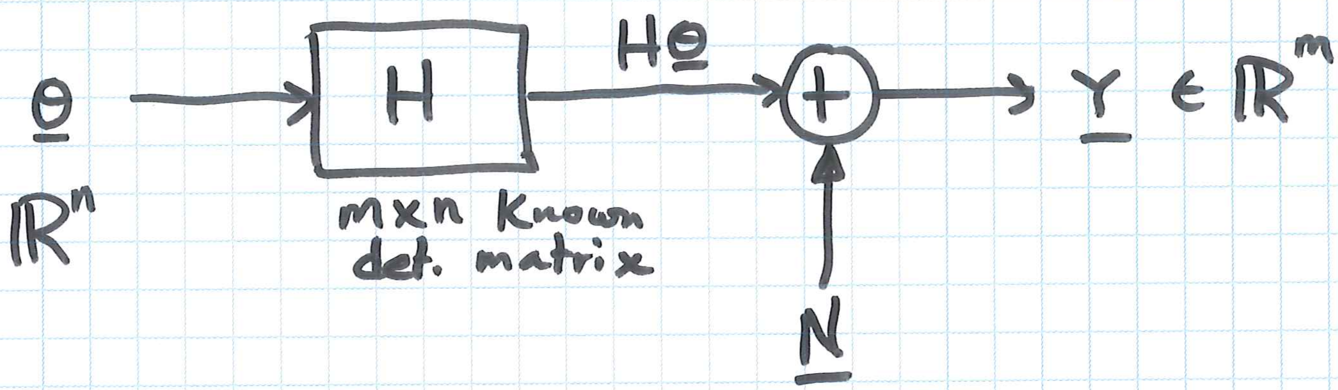
$$\underset{B \in \mathcal{B}}{P_X(B)} = \int_B f_X(x)\,dx.$$

Models for $\{P_\theta : \theta \in \Lambda\}$



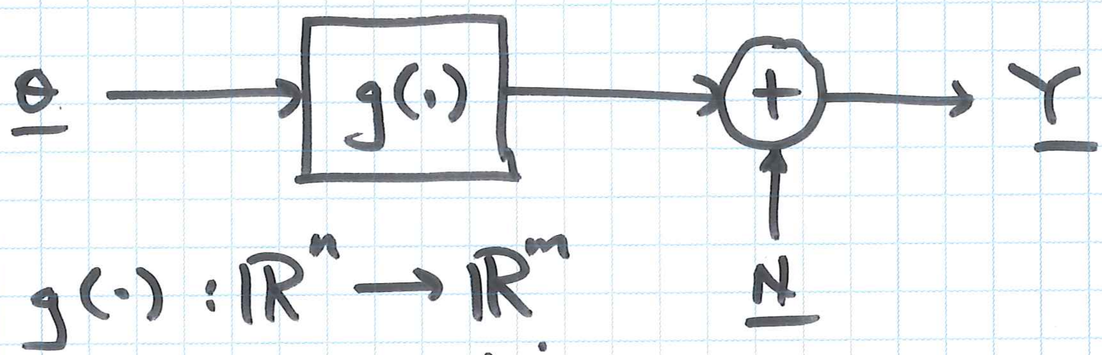
Assumptions here give diff. model types.

Linear Additive Noise Model



$$\underline{Y} = H\underline{\Theta} + N$$

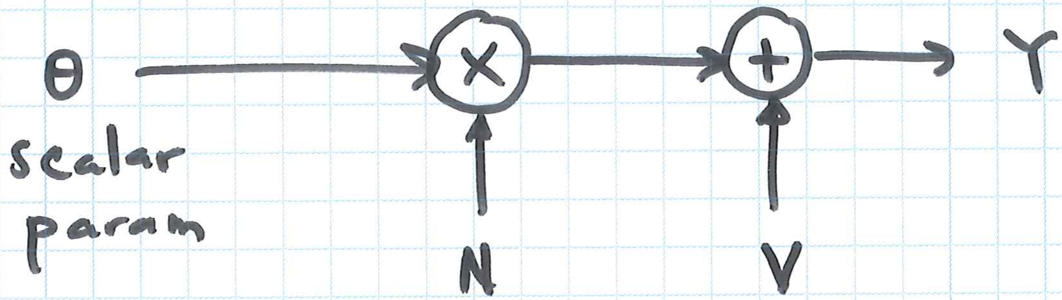
Non-linear Additive Noise Model



$g(\cdot) : \mathbb{R}^n \longrightarrow \mathbb{R}^m$

known, nonlinear

Multiplicative Noise Model



$$Y = N\theta + V$$

The Unknown Parameter θ

① Classical Model $\{f_\theta(y) : \theta \in \Lambda\}$ pdfs for Y.

No further structure for Θ.

② Bayesian Model

Here the model pdfs are thought to be conditional pdfs

$$f_\theta(y) = f_{Y|\Theta}(y|\theta)$$

Θ is a r.v.

Also have a prior $f_\Theta(\theta)$.

$$\Rightarrow f_{Y,\Theta}(y|\theta) = f_{Y|\Theta}(y|\theta)\, f_\Theta(\theta).$$

Definition of Estimator

Is a mapping of values of $Y=y$ into values of param.

$$\hat{\theta}(y) : \Gamma \rightarrow \Lambda$$

interpete two ways!

1) $\hat{\theta}(y)$ is an estimate
2) $\hat{\theta}(Y)$ is a estimator

What is "best", "good", "optimal"

come down to MSE
or
Prob. of Error.

4.1 Real-Valued Gaussian Random Variables

1. A real-valued random variable X is said to be Normal or Gaussian with mean μ and variance σ^2 if its pdf is

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}$$

   for $x \in \mathcal{R}$.

2. This fact is often noted by $X \sim \mathcal{N}(\mu, \sigma^2)$.

3. Note that if $X \sim \mathcal{N}(\mu, \sigma^2)$, then

$$\frac{X-\mu}{\sigma} \sim \mathcal{N}(0,1).$$

4. For such rvs all finite order moments exist. In fact, if $X \sim \mathcal{N}(0,1)$, then

$$\begin{aligned} \mathrm{E}\{X^{2k}\} &= 1 \cdot 3 \cdot 5 \cdots (2k-1) \\ \mathrm{E}\{X^{2k-1}\} &= 0 \end{aligned}$$

   for $k = 1, 2, 3, \ldots$.

AMPAD™

4.2 Real-Valued Gaussian Random Vectors

1. A real vector valued random variable of dimension n

$$\boldsymbol{X} = \begin{bmatrix} X_1 & X_2 & \cdots & X_n \end{bmatrix}^T$$

is said to be multivariate normal if its joint pdf[4] is of the form:

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{(2\pi)^{n/2}(\det\,\Sigma)^{1/2}} \exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\}$$

for $\boldsymbol{x} \in \mathcal{R}^n$.

2. A shorthand notation is $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$.

3. If $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$ then can be shown:

   (a) Mean vector: $\mathrm{E}\{\boldsymbol{X}\} = \boldsymbol{\mu}$.

   (b) Covariance matrix: $\mathrm{E}\{(\boldsymbol{X}-\boldsymbol{\mu})(\boldsymbol{X}-\boldsymbol{\mu})^T\} = \Sigma$.

   (c) Correlation matrix: $\boldsymbol{R} = \mathrm{E}\{\boldsymbol{X}\boldsymbol{X}^T\} = \Sigma + \boldsymbol{\mu}\boldsymbol{\mu}^T$.

[4] Here we assume the matrix Σ is non-singular. Gaussian random vectors can be defined even when the matrix is singular, as we will discuss later.

Back to Estimation with Gaussian random vectors. Let Z be multivariate Normal

$$Z \sim \eta(\mu, \Sigma)$$

and suppose $\Sigma > 0$ and that the vector Z is partitioned into two random vectors X, Y according to

$$Z = \begin{bmatrix} X \\ Y \end{bmatrix} \qquad \mu = \begin{bmatrix} \mu_X \\ \mu_Y \end{bmatrix} \qquad \Sigma = \begin{bmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{XY}^T & \Sigma_{YY} \end{bmatrix}$$

Then we have already argued that

$$X \sim \eta(\mu_X\, \Sigma_{XX}) \qquad Y \sim \eta(\mu_Y\, \Sigma_{YY})$$

$$X \perp\!\!\!\perp Y \iff \Sigma_{XY} = 0$$

<u>Claim</u> Subject to above hypotheses, the conditional distribution of X given Y is multivariate Normal with <u>mean</u>

$$\mu_X + \Sigma_{XY} \Sigma_{YY}^{-1} (y - \mu_Y)$$

and <u>covariance</u>

$$\Sigma_{XX} - \Sigma_{XY} \Sigma_{YY}^{-1} \Sigma_{XY}^T$$

Condition mean estimator is best among all poss. Borel functions of observations

We try to show this using the direct calculation

$$f_{X|Y}(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)} = \frac{f_Z(z)}{f_Y(y)}$$

$$= \frac{(2\pi)^{-n/2} |\Sigma|^{-1/2} \exp\left\{-\frac{1}{2}(z-\mu)^T \Sigma^{-1} (z-\mu)\right\}}{(2\pi)^{-(n-p)/2} |\Sigma_{YY}|^{-1/2} \exp\left\{-\frac{1}{2}(y-\mu_Y)^T \Sigma_{YY}^{-1} (y-\mu_Y)\right\}}$$

where $n = \dim Z$
$p = \dim X$
$n-p = \dim Y$