Introduction to (Bayesian) Estimation
MAE 5020

Models

Oklahoma State University

August 16, 2023

Introduction to the generic linear model

- To estimate unknown quantitie based on measurements and other given information, begin with a model representation
- Explicit relationship between the unknowns and the measurements
- Linear vs. nonlinear relationship
- Fundamental estimation problems in a generic linear model

$$Z(k) = H(k)\theta + V(k)$$

$\theta \in \mathbb{R}^{n\times 1}$:
$Z(k) \in \mathbb{R}^{N\times 1}$:
$H(k) \in \mathbb{R}^{N\times n}$:
$V(k) \in \mathbb{R}^{N\times 1}$:

Categories

A. θ is deterministic
   1. $H(k)$ is deterministic
   2. $H(k)$ is random
      a) $H(k)$ and $V(k)$ are statistically independent
      b) $H(k)$ and $V(k)$ are statistically dependent

B. θ is random
   1. $H(k)$ is deterministic
   2. $H(k)$ is random
      a) $H(k)$ and $V(k)$ are statistically independent
      b) $H(k)$ and $V(k)$ are statistically dependent

In the book, there are many examples. We will discuss a few that are important and instrumental.

Example 2-3: Function approximation

- Given $(x_1, f(x_1)), (x_2, f(x_2)), \cdots, (x_N, f(x_N))$, where x_i's are unknown but $f(\cdot)$ is unknown.
- Approxate $f(\cdot)$ in a linear form

$$f(x) \approx \hat{f}(x) \triangleq \sum_{j=1}^{n} \theta_j \phi_j(x)$$

- $\phi_j(x)$:
- Examples:

The corresponding linear model

- Suppose $f_m(x_i) = f(x_i) + e(x_i)$.

The corresponding linear model

- Suppose $f_m(x_i) = f(x_i) + e(x_i)$.
- Rewrite it in a vector form with $\hat{f}(\cdot)$.

- Final form in the generic linear model:

- Category:

Example 2-1: Finite impulse response (FIR)

FIR Model (a.k.a Moving average model):

$$\underbrace{y(k)}_{output} = \sum_{i=1}^{n} h(i)\,\underbrace{u(k-i)}_{input}$$

$$z(k) = y(k) + v(k)$$

- Moving average coefficients $h(i)$'s are unknown
- "signal-plus-noise" model in block diagram:

The corresponding linear model

- Rewrite it in a vector form
- The generic linear model
- Category:

Example 2-2: Finite difference equation

Autoregressive (AR) model:

$$\underbrace{y(k)}_{measurement} = -\alpha_1 y(k-1) - \alpha_2 y(k-2) - \cdots - \alpha_n y(k-n) + u(k-1)$$

- Rewrite it in a vector form:

- The generic linear model:

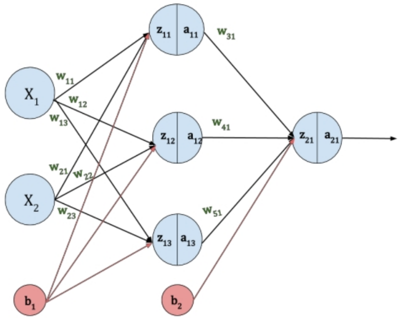
- Category

Example 2-5: Nonlinear model

- Most of the developed techniques can be applied to nonlinear models (after linearization)
- Suppose $z(k) = f(\theta, k) + v(k)$, $k = 1, \cdots, N$.
- f : nonlinear function in θ, known explicitly.
- Nominal θ: θ^*, nominal $z^*(k) =$
- Linearization: define $\delta z(k) =$, $\delta\theta =$
- Linearized equation:

The corresponding linear form

Example (new): Neural network as a nonlinear model



A simple NN model [a].

First layer:

$$\begin{pmatrix} z_{11} \\ z_{12} \\ z_{13} \end{pmatrix} = \begin{pmatrix} w_{11} & w_{21} \\ w_{12} & w_{22} \\ w_{13} & w_{23} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} b_1 \\ b_1 \\ b_1 \end{pmatrix}$$

$$Z^{[1]} = W^{[1]}X + b^{[1]}$$

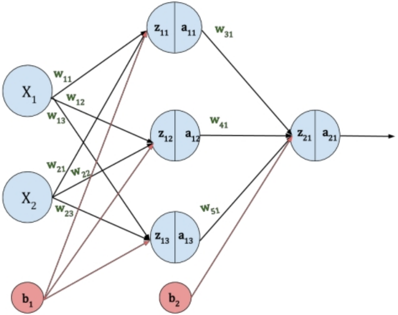
Activation function:

$$\begin{pmatrix} a_{11} \\ a_{12} \\ a_{13} \end{pmatrix} = \begin{pmatrix} \sigma(a_{11}) \\ \sigma(a_{12}) \\ \sigma(a_{13}) \end{pmatrix}$$

$$A^{[1]} = \sigma(Z^{[1]})$$

[a] https://studymachinelearning.com/mathematics-behind-the-neural-network/

Example (new): Neural network as a nonlinear model



A simple NN model.

Second layer:

$$\begin{pmatrix} z_{21} \end{pmatrix} = \begin{pmatrix} w_{31} & w_{41} & w_{51} \end{pmatrix} \begin{pmatrix} a_{11} \\ a_{12} \\ a_{13} \end{pmatrix} + b_2$$

$$Z^{[2]} = W^{[2]} A^{[1]} + b^{[2]}$$

Activation function:
$a_{21} = \sigma(z_{21})$, $A^{[2]} = \sigma(Z^{[2]})$.
Putting everything together:

$$A^{[2]} = \sigma(W^{[2]}\sigma(W^{[1]}X + b^{[1]}) + b^{[2]})$$

$$\triangleq f(\theta, X)$$

Example 2-4: dynamic systems and state estimation

- Dynamical systems are ubiquitous in control/robotics/communication, etc.
- Often we need to estimate the state vector in a dynamical system for controls, monitoring, etc.
- Typically we cannot measure all the state variables, e.g., position but not velocity, etc.
- State estimation: estimate the entire state vector given a limited collection of noisy measurements

Formulation

$$x(k+1) = \Phi x(k) + \gamma u(k)$$
$$z(k) = h^T x(k) + v(k)$$

Collect N measurements:

Find a common θ by using the state equation. First, state solution:

Second, fix to a particular time k_1:

The general linear form

- Express measurements in terms of $x(k_1)$
- The linear form

Observations