Tutorial 6, Mar 1, 2024

Gaussian Parameter Estimation

- Let $\boldsymbol{W} = \begin{bmatrix} X_1 & \dots & X_n & Y_1 & \dots & Y_n \end{bmatrix}^T \sim \mathcal{N}(\boldsymbol{\mu}_W, \boldsymbol{\Sigma}_W)$; suppose we observe $\boldsymbol{Y} = \begin{bmatrix} Y_1 & \dots & Y_n \end{bmatrix}^T$, we wish to estimate $\boldsymbol{X} = \begin{bmatrix} X_1 & \dots & X_n \end{bmatrix}^T$
- Claim: $f(\boldsymbol{x}|\boldsymbol{y}) \sim \mathcal{N}(\boldsymbol{\mu}_{X|Y}, \boldsymbol{\Sigma}_{X|Y})$; if this is true, since the max of a Gaussian is at its mean, the MAP estimate of $\boldsymbol{x}$ given $\boldsymbol{y}$ is simply the mean of the Gaussian, i.e. $\boldsymbol{\mu}_{X|Y}$
  - Proof in course notes
  - $\boldsymbol{\mu}_{X|Y} = \boldsymbol{\mu}_X + \boldsymbol{\Sigma}_{XY}\boldsymbol{\Sigma}_{YY}(\boldsymbol{y} - \boldsymbol{\mu}_Y)$
  - $\boldsymbol{\Sigma}_{X|Y} = \boldsymbol{\Sigma}_{XX} - \boldsymbol{\Sigma}_{XY}\boldsymbol{\Sigma}_{YY}^{-1}\boldsymbol{\Sigma}_{YX}$
  - Where $\boldsymbol{\Sigma}_W = \begin{bmatrix} \boldsymbol{\Sigma}_{\boldsymbol{XX}} & \boldsymbol{\Sigma}_{\boldsymbol{XY}} \\ \boldsymbol{\Sigma}_{\boldsymbol{YX}} & \boldsymbol{\Sigma}_{\boldsymbol{YY}} \end{bmatrix}$
- Example: let $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}_X, \boldsymbol{\Sigma}_X), \boldsymbol{Z} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_Z), \boldsymbol{A} \in \mathbb{R}^{n\times n}, \boldsymbol{b} \in \mathbb{R}^n$ and consider $\boldsymbol{Y} = \boldsymbol{AX} + \boldsymbol{Z} + \boldsymbol{b}$; given an observation of $\boldsymbol{Y}$, we wish to estimate $\boldsymbol{X}$
  - Let $\boldsymbol{W} = \begin{bmatrix} \boldsymbol{X} \\ \boldsymbol{Y} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X} \\ \boldsymbol{AX} + \boldsymbol{Z} + \boldsymbol{b} \end{bmatrix} = \begin{bmatrix} \boldsymbol{1}_{n\times n} & \boldsymbol{0} \\ \boldsymbol{A} & \boldsymbol{1}_{n\times n} \end{bmatrix} \begin{bmatrix} \boldsymbol{X} \\ \boldsymbol{Z} \end{bmatrix} + \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{b} \end{bmatrix}$
  - Notice that $\begin{bmatrix} \boldsymbol{X} \\ \boldsymbol{Z} \end{bmatrix}$ is Gaussian, and $\begin{bmatrix} \boldsymbol{X} \\ \boldsymbol{Z} \end{bmatrix} \to \begin{bmatrix} \boldsymbol{X} \\ \boldsymbol{Y} \end{bmatrix}$ by a linear transformation, so it is also Gaussian
  - We can use the parameter estimation formula shown above once we find the mean and covariance matrix of $\boldsymbol{W} = \begin{bmatrix} \boldsymbol{X} \\ \boldsymbol{Y} \end{bmatrix}$
- Example: suppose $X, W_1, W_2 \sim \mathcal{N}(0,1)$ and independent; let $Y_1 = X + W_1, Y_2 = X + W_2, Y_3 = W_1 + W_2$; find the MAP estimate of X given Y_1, Y_2, Y_3
  - We get two noisy observation of X and an observation of the noise
  - $\boldsymbol{W} = \begin{bmatrix} X \\ Y_1 \\ X_2 \\ Y_3 \end{bmatrix} = \begin{bmatrix} X \\ X + W_1 \\ X + W_2 \\ W_1 + W_2 \end{bmatrix}$
    * Note this is Gaussian since all the terms are Gaussian, so their sum will also be
  - $\boldsymbol{\Sigma}_W = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 0 & 1 & 1 & 2 \end{bmatrix}$
    * For diagonal entries: variance of X is 1; variance of the sum of two unit variance variables is 2
    * For off-diagonal entries use the bilinearity of covariance and note that X, W_1, W_2 are all independent
  - Partition the above into sub-matrices for the covariances and use the formulas