ECE 275A: Parameter Estimation I
The Multivariate Gaussian Distribution

Florian Meyer

Electrical and Computer Engineering Department
University of California San Diego

UC San Diego
JACOBS SCHOOL OF ENGINEERING

Multivariate Gaussian Distribution

- Suppose the observation $\mathbf{y} \in \mathbb{R}^m$ and the unknown parameter $\boldsymbol{\theta} \in \mathbb{R}^p$ are jointly Gaussian
- Let $\mathbf{z} = \begin{bmatrix} \boldsymbol{\theta} \\ \mathbf{y} \end{bmatrix} \in \mathbb{R}^{p+m}$, then

$$\mathbf{z} \sim \mathcal{N}(\boldsymbol{m_z}, \boldsymbol{C_{zz}}) \text{ with } \boldsymbol{m_z} = \begin{bmatrix} \boldsymbol{m_\theta} \\ \boldsymbol{m_y} \end{bmatrix}, \boldsymbol{C_{zz}} = \begin{bmatrix} \boldsymbol{C_{\theta\theta}} & \boldsymbol{C_{\theta y}} \\ \boldsymbol{C_{y\theta}} & \boldsymbol{C_{yy}} \end{bmatrix}, \boldsymbol{C_{\theta y}} = \boldsymbol{C_{y\theta}}^\top$$

- Important properties of multivariate Gaussian distributions:
  1. Conditional distributions $p(\boldsymbol{\theta}|\boldsymbol{y}), p(\boldsymbol{y}|\boldsymbol{\theta})$ are Gaussian
  2. Marginal distributions $p(\boldsymbol{y}), p(\boldsymbol{\theta})$ are Gaussian

Multivariate Gaussian Distribution

- $\boldsymbol{C}_{\boldsymbol{zz}}$ can be decomposed as

$$\boldsymbol{C}_{\boldsymbol{zz}} = \begin{bmatrix} \boldsymbol{C}_{\theta\theta} & \boldsymbol{C}_{\theta \boldsymbol{y}} \\ \boldsymbol{C}_{\boldsymbol{y}\theta} & \boldsymbol{C}_{\boldsymbol{yy}} \end{bmatrix} = \underbrace{\begin{bmatrix} \boldsymbol{I} & \boldsymbol{C}_{\theta \boldsymbol{y}} \boldsymbol{C}_{\boldsymbol{yy}}^{-1} \\ \boldsymbol{0} & \boldsymbol{I} \end{bmatrix}}_{\boldsymbol{A}} \underbrace{\begin{bmatrix} \boldsymbol{C}_{\theta\theta|\boldsymbol{y}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{C}_{\boldsymbol{yy}} \end{bmatrix}}_{\boldsymbol{B}} \underbrace{\begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{C}_{\boldsymbol{yy}}^{-1} \boldsymbol{C}_{\boldsymbol{y}\theta} & \boldsymbol{I} \end{bmatrix}}_{\boldsymbol{A}^{\mathsf{T}}}$$

  where $\boldsymbol{C}_{\theta\theta|\boldsymbol{y}} = \boldsymbol{C}_{\theta\theta} - \boldsymbol{C}_{\theta \boldsymbol{y}} \boldsymbol{C}_{\boldsymbol{yy}}^{-1} \boldsymbol{C}_{\boldsymbol{y}\theta}$ and it is also known as the Schur complement of $\boldsymbol{C}_{\boldsymbol{yy}}$

- Furthermore, we have $|\boldsymbol{C}_{\boldsymbol{zz}}| = |\boldsymbol{C}_{\theta\theta|\boldsymbol{y}}| \cdot |\boldsymbol{C}_{\boldsymbol{yy}}|$ and

$$\boldsymbol{C}_{\boldsymbol{zz}}^{-1} = \left(\boldsymbol{A}^{-1}\right)^{\mathsf{T}} \boldsymbol{B}^{-1} \boldsymbol{A}^{-1} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ -\boldsymbol{C}_{\boldsymbol{yy}}^{-1} \boldsymbol{C}_{\boldsymbol{y}\theta} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{C}_{\theta\theta|\boldsymbol{y}}^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{C}_{\boldsymbol{yy}}^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{I} & -\boldsymbol{C}_{\theta \boldsymbol{y}} \boldsymbol{C}_{\boldsymbol{yy}}^{-1} \\ \boldsymbol{0} & \boldsymbol{I} \end{bmatrix}$$

Multivariate Gaussian Distribution

- Note that

$$
\begin{aligned}
-\frac{1}{2}\left\|\boldsymbol{z}-\boldsymbol{m}_{\boldsymbol{z}}\right\|_{\boldsymbol{C}_{\boldsymbol{zz}}^{-1}}^{2} &= -\frac{1}{2}\left\| \begin{bmatrix} \theta-\boldsymbol{m}_{\theta} \\ \boldsymbol{y}-\boldsymbol{m}_{\boldsymbol{y}} \end{bmatrix} \right\|_{\boldsymbol{C}_{\boldsymbol{zz}}^{-1}}^{2} \\
&= -\frac{1}{2}\begin{bmatrix} \theta-\boldsymbol{m}_{\theta} \\ \boldsymbol{y}-\boldsymbol{m}_{\boldsymbol{y}} \end{bmatrix}^{\mathsf{T}} \left(\boldsymbol{A}^{-1}\right)^{\mathsf{T}} \boldsymbol{B}^{-1}\boldsymbol{A}^{-1} \begin{bmatrix} \theta-\boldsymbol{m}_{\theta} \\ \boldsymbol{y}-\boldsymbol{m}_{\boldsymbol{y}} \end{bmatrix} \\
&= -\frac{1}{2}\begin{bmatrix} \theta-\boldsymbol{m}_{\theta|\boldsymbol{y}} \\ \boldsymbol{y}-\boldsymbol{m}_{\boldsymbol{y}} \end{bmatrix}^{\mathsf{T}} \boldsymbol{B}^{-1} \begin{bmatrix} \theta-\boldsymbol{m}_{\theta|\boldsymbol{y}} \\ \boldsymbol{y}-\boldsymbol{m}_{\boldsymbol{y}} \end{bmatrix} \\
&= -\frac{1}{2}\begin{bmatrix} \theta-\boldsymbol{m}_{\theta|\boldsymbol{y}} \\ \boldsymbol{y}-\boldsymbol{m}_{\boldsymbol{y}} \end{bmatrix}^{\mathsf{T}} \begin{bmatrix} \boldsymbol{C}_{\theta\theta|\boldsymbol{y}}^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{C}_{\boldsymbol{yy}}^{-1} \end{bmatrix} \begin{bmatrix} \theta-\boldsymbol{m}_{\theta|\boldsymbol{y}} \\ \boldsymbol{y}-\boldsymbol{m}_{\boldsymbol{y}} \end{bmatrix} \\
&= -\frac{1}{2}\left\|\theta-\boldsymbol{m}_{\theta|\boldsymbol{y}}\right\|_{\boldsymbol{C}_{\theta\theta|\boldsymbol{y}}^{-1}}^{2} - \frac{1}{2}\left\|\boldsymbol{y}-\boldsymbol{m}_{\boldsymbol{y}}\right\|_{\boldsymbol{C}_{\boldsymbol{yy}}^{-1}}^{2}
\end{aligned}
$$

where $\boldsymbol{m}_{\theta|\boldsymbol{y}} = \boldsymbol{m}_{\theta} + \boldsymbol{C}_{\theta\boldsymbol{y}}\boldsymbol{C}_{\boldsymbol{yy}}^{-1}(\boldsymbol{y}-\boldsymbol{m}_{\boldsymbol{y}})$

Multivariate Gaussian Distribution

$$p(\boldsymbol{\theta}, \boldsymbol{y}) = p(\boldsymbol{z}) = \frac{\exp\Big(-\frac{1}{2}\big\|\boldsymbol{z} - \boldsymbol{m_z}\big\|^2_{\boldsymbol{C}_{\boldsymbol{zz}}^{-1}}\Big)}{(2\pi)^{\frac{p+m}{2}}\big|\boldsymbol{C_{zz}}\big|^{\frac{1}{2}}}$$

$$= \frac{\exp\Big(-\frac{1}{2}\big\|\boldsymbol{\theta} - \boldsymbol{m}_{\theta|\boldsymbol{y}}\big\|^2_{\boldsymbol{C}_{\boldsymbol{\theta\theta}|\boldsymbol{y}}^{-1}}\Big)}{(2\pi)^{\frac{p}{2}}\big|\boldsymbol{C}_{\boldsymbol{\theta\theta}|\boldsymbol{y}}\big|^{\frac{1}{2}}} \cdot \frac{\exp\Big(-\frac{1}{2}\big\|\boldsymbol{y} - \boldsymbol{m_y}\big\|^2_{\boldsymbol{C}_{\boldsymbol{yy}}^{-1}}\Big)}{(2\pi)^{\frac{m}{2}}\big|\boldsymbol{C_{yy}}\big|^{\frac{1}{2}}}$$

$$p(\boldsymbol{y}) = \int p(\boldsymbol{\theta}, \boldsymbol{y}) d\boldsymbol{\theta} = \frac{\exp\Big(-\frac{1}{2}\big\|\boldsymbol{y} - \boldsymbol{m_y}\big\|^2_{\boldsymbol{C}_{\boldsymbol{yy}}^{-1}}\Big)}{(2\pi)^{\frac{m}{2}}\big|\boldsymbol{C_{yy}}\big|^{\frac{1}{2}}} \text{ is Gaussian}$$

$$p(\boldsymbol{\theta}|\boldsymbol{y}) = \frac{p(\boldsymbol{\theta}, \boldsymbol{y})}{p(\boldsymbol{y})} = \frac{\exp\Big(-\frac{1}{2}\big\|\boldsymbol{\theta} - \boldsymbol{m}_{\theta|\boldsymbol{y}}\big\|^2_{\boldsymbol{C}_{\boldsymbol{\theta\theta}|\boldsymbol{y}}^{-1}}\Big)}{(2\pi)^{\frac{p}{2}}\big|\boldsymbol{C}_{\boldsymbol{\theta\theta}|\boldsymbol{y}}\big|^{\frac{1}{2}}} \text{ is Gaussian}$$