

430.714

Estimation Theory

Prof. Songhwai Oh
ECE, SNU

Multivariate Gaussian

Theorem 10.2 (Conditional PDF of Multivariate Gaussian) *If $\mathbf{x}$ and $\mathbf{y}$ are jointly Gaussian, where $\mathbf{x}$ is $k \times 1$ and $\mathbf{y}$ is $l \times 1$, with mean vector $[E(\mathbf{x})^T E(\mathbf{y})^T]^T$ and partitioned covariance matrix*

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_{xx} & \mathbf{C}_{xy} \\ \mathbf{C}_{yx} & \mathbf{C}_{yy} \end{bmatrix} = \begin{bmatrix} k \times k & k \times l \\ l \times k & l \times l \end{bmatrix} \quad (10.23)$$

so that

$$p(\mathbf{x}, \mathbf{y}) = \frac{1}{(2\pi)^{\frac{k+l}{2}} \det^{\frac{1}{2}}(\mathbf{C})} \exp \left[-\frac{1}{2} \left(\begin{bmatrix} \mathbf{x} - E(\mathbf{x}) \\ \mathbf{y} - E(\mathbf{y}) \end{bmatrix} \right)^T \mathbf{C}^{-1} \left(\begin{bmatrix} \mathbf{x} - E(\mathbf{x}) \\ \mathbf{y} - E(\mathbf{y}) \end{bmatrix} \right) \right],$$

then the conditional PDF $p(\mathbf{y}|\mathbf{x})$ is also Gaussian and

$$E(\mathbf{y}|\mathbf{x}) = E(\mathbf{y}) + \mathbf{C}_{yx} \mathbf{C}_{xx}^{-1} (\mathbf{x} - E(\mathbf{x})) \quad (10.24)$$

$$\mathbf{C}_{y|x} = \mathbf{C}_{yy} - \mathbf{C}_{yx} \mathbf{C}_{xx}^{-1} \mathbf{C}_{xy}. \quad (10.25)$$

Conditional PDF of Multivariate Gaussian: Bivariate Case

$$\mathbf{y}|\mathbf{x} \sim \mathcal{N}(\mathbb{E}(\mathbf{y}|\mathbf{x}), \mathbf{C}_{y|x})$$

$$\begin{aligned} E(\mathbf{y}|\mathbf{x}) &= E(\mathbf{y}) + \mathbf{C}_{yx}\mathbf{C}_{xx}^{-1}(\mathbf{x} - E(\mathbf{x})) \\ \mathbf{C}_{y|x} &= \mathbf{C}_{yy} - \mathbf{C}_{yx}\mathbf{C}_{xx}^{-1}\mathbf{C}_{xy}. \end{aligned}$$

Bivariate Gaussian r.v. $[x, y]^T$

$$\mathbf{C} = \begin{bmatrix} \text{var}(x) & \text{cov}(x, y) \\ \text{cov}(y, x) & \text{var}(y) \end{bmatrix}$$

$$E(y|x) = E(y) + \frac{\text{cov}(x, y)}{\text{var}(x)} (x - E(x))$$

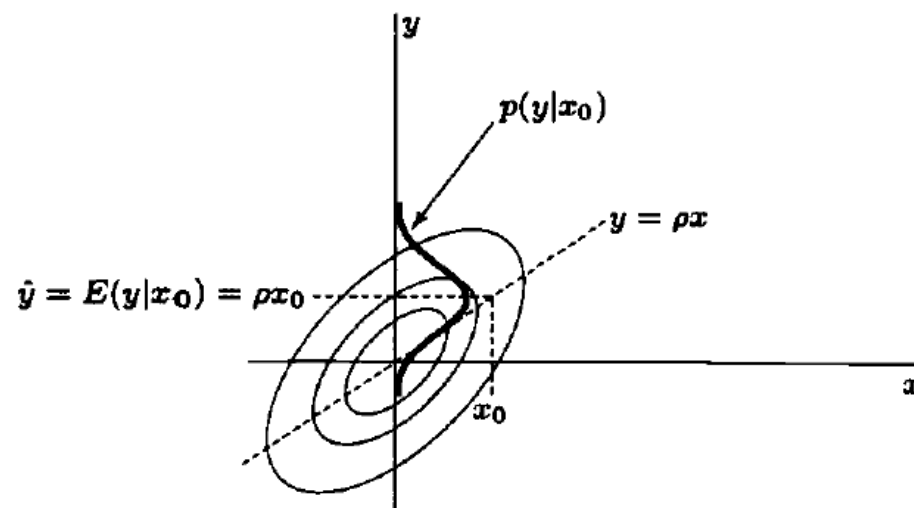
$$\text{var}(y|x) = \text{var}(y) - \frac{\text{cov}^2(x, y)}{\text{var}(x)}.$$

Correlation coefficient: $\rho = \frac{\text{cov}(x, y)}{\sqrt{\text{var}(x)\text{var}(y)}}$

$$\frac{\hat{y} - E(y)}{\sqrt{\text{var}(y)}} = \frac{\text{cov}(x, y)}{\sqrt{\text{var}(x)\text{var}(y)}} \frac{x - E(x)}{\sqrt{\text{var}(x)}}$$

$$\hat{y}_n = \rho x_n$$

$$\begin{aligned} \text{var}(y|x) &= \text{var}(y) \left[1 - \frac{\text{cov}^2(x, y)}{\text{var}(x)\text{var}(y)} \right] \\ &= \text{var}(y)(1 - \rho^2) \end{aligned}$$



Conditional PDF of Multivariate Gaussian: Derivation

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

$$= \frac{\frac{1}{(2\pi)^{\frac{k+l}{2}} \det^{\frac{1}{2}}(\mathbf{C})} \exp \left[-\frac{1}{2} \left(\begin{bmatrix} \mathbf{x} - E(\mathbf{x}) \\ \mathbf{y} - E(\mathbf{y}) \end{bmatrix} \right)^T \mathbf{C}^{-1} \left(\begin{bmatrix} \mathbf{x} - E(\mathbf{x}) \\ \mathbf{y} - E(\mathbf{y}) \end{bmatrix} \right) \right]}{\frac{1}{(2\pi)^{\frac{k}{2}} \det^{\frac{1}{2}}(\mathbf{C}_{xx})} \exp \left[-\frac{1}{2} (\mathbf{x} - E(\mathbf{x}))^T \mathbf{C}_{xx}^{-1} (\mathbf{x} - E(\mathbf{x})) \right]}$$

$$\begin{bmatrix} I & 0 \\ -A_{21}A_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} I & -A_{11}^{-1}A_{12} \\ 0 & I \end{bmatrix} = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{bmatrix}$$

$$\det \left(\begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \right) = \det(\mathbf{A}_{11}) \det(\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})$$

Schur complement
(of A wrt A_{11})

$$\det(\mathbf{C}) = \det(\mathbf{C}_{xx}) \det(\mathbf{C}_{yy} - \mathbf{C}_{yx}\mathbf{C}_{xx}^{-1}\mathbf{C}_{xy}) \quad \frac{\det(\mathbf{C})}{\det(\mathbf{C}_{xx})} = \det(\mathbf{C}_{yy} - \mathbf{C}_{yx}\mathbf{C}_{xx}^{-1}\mathbf{C}_{xy})$$

$$p(\mathbf{y}|\mathbf{x}) = \frac{1}{(2\pi)^{\frac{l}{2}} \det^{\frac{1}{2}}(\mathbf{C}_{yy} - \mathbf{C}_{yx}\mathbf{C}_{xx}^{-1}\mathbf{C}_{xy})} \exp \left(-\frac{1}{2} Q \right)$$

$$Q = \begin{bmatrix} \mathbf{x} - E(\mathbf{x}) \\ \mathbf{y} - E(\mathbf{y}) \end{bmatrix}^T \mathbf{C}^{-1} \begin{bmatrix} \mathbf{x} - E(\mathbf{x}) \\ \mathbf{y} - E(\mathbf{y}) \end{bmatrix} - (\mathbf{x} - E(\mathbf{x}))^T \mathbf{C}_{xx}^{-1} (\mathbf{x} - E(\mathbf{x}))$$

Conditional PDF of Multivariate Gaussian: Derivation

$$\begin{bmatrix} I & 0 \\ -A_{21}A_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} I & -A_{11}^{-1}A_{12} \\ 0 & I \end{bmatrix} = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{bmatrix} \quad \begin{aligned} XYZ &= W \\ Y^{-1} &= ZW^{-1}X \end{aligned}$$

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_{xx} & \mathbf{C}_{xy} \\ \mathbf{C}_{yx} & \mathbf{C}_{yy} \end{bmatrix} \quad \mathbf{C}^{-1} = \begin{bmatrix} \mathbf{I} & -\mathbf{C}_{xx}^{-1}\mathbf{C}_{xy} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{C}_{xx}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{C}_{yx}\mathbf{C}_{xx}^{-1} & \mathbf{I} \end{bmatrix} \quad \mathbf{B} = \mathbf{C}_{yy} - \mathbf{C}_{yx}\mathbf{C}_{xx}^{-1}\mathbf{C}_{xy}$$

$$Q = \begin{bmatrix} \mathbf{x} - E(\mathbf{x}) \\ \mathbf{y} - E(\mathbf{y}) \end{bmatrix}^T \mathbf{C}^{-1} \begin{bmatrix} \mathbf{x} - E(\mathbf{x}) \\ \mathbf{y} - E(\mathbf{y}) \end{bmatrix} - (\mathbf{x} - E(\mathbf{x}))^T \mathbf{C}_{xx}^{-1} (\mathbf{x} - E(\mathbf{x}))$$

$$Q = \begin{bmatrix} \tilde{\mathbf{x}} \\ \tilde{\mathbf{y}} \end{bmatrix}^T \begin{bmatrix} \mathbf{I} & -\mathbf{C}_{xx}^{-1}\mathbf{C}_{xy} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{C}_{xx}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{C}_{yx}\mathbf{C}_{xx}^{-1} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{x}} \\ \tilde{\mathbf{y}} \end{bmatrix} \quad \begin{aligned} \tilde{\mathbf{x}} &= \mathbf{x} - E(\mathbf{x}) \\ \tilde{\mathbf{y}} &= \mathbf{y} - E(\mathbf{y}) \end{aligned}$$

$$\begin{aligned} & -\tilde{\mathbf{x}}^T \mathbf{C}_{xx}^{-1} \tilde{\mathbf{x}} \\ &= \begin{bmatrix} \tilde{\mathbf{x}} \\ \tilde{\mathbf{y}} - \mathbf{C}_{yx}\mathbf{C}_{xx}^{-1}\tilde{\mathbf{x}} \end{bmatrix}^T \begin{bmatrix} \mathbf{C}_{xx}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}^{-1} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{x}} \\ \tilde{\mathbf{y}} - \mathbf{C}_{yx}\mathbf{C}_{xx}^{-1}\tilde{\mathbf{x}} \end{bmatrix} - \tilde{\mathbf{x}}^T \mathbf{C}_{xx}^{-1} \tilde{\mathbf{x}} \\ &= (\tilde{\mathbf{y}} - \mathbf{C}_{yx}\mathbf{C}_{xx}^{-1}\tilde{\mathbf{x}})^T \mathbf{B}^{-1} (\tilde{\mathbf{y}} - \mathbf{C}_{yx}\mathbf{C}_{xx}^{-1}\tilde{\mathbf{x}}) \\ &= [\mathbf{y} - (E(\mathbf{y}) + \mathbf{C}_{yx}\mathbf{C}_{xx}^{-1}(\mathbf{x} - E(\mathbf{x})))]^T \\ & \quad [\mathbf{C}_{yy} - \mathbf{C}_{yx}\mathbf{C}_{xx}^{-1}\mathbf{C}_{xy}]^{-1} [\mathbf{y} - (E(\mathbf{y}) + \mathbf{C}_{yx}\mathbf{C}_{xx}^{-1}(\mathbf{x} - E(\mathbf{x})))] \end{aligned}$$



$$\begin{aligned} \mathbf{y}|\mathbf{x} &\sim \mathcal{N}(\mathbb{E}(\mathbf{y}|\mathbf{x}), \mathbf{C}_{y|x}) \\ \mathbb{E}(\mathbf{y}|\mathbf{x}) &= E(\mathbf{y}) + \mathbf{C}_{yx}\mathbf{C}_{xx}^{-1}(\mathbf{x} - E(\mathbf{x})) \\ \mathbf{C}_{y|x} &= \mathbf{C}_{yy} - \mathbf{C}_{yx}\mathbf{C}_{xx}^{-1}\mathbf{C}_{xy} \end{aligned}$$

Matrix Inversion Lemma

Suppose A_{11} and A_{22} are invertible and let $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$

$$\begin{bmatrix} I & 0 \\ -A_{21}A_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} I & -A_{11}^{-1}A_{12} \\ 0 & I \end{bmatrix} = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{bmatrix} \quad \begin{array}{l} XYZ = W \\ Y^{-1} = ZW^{-1}X \end{array}$$

$$\begin{aligned} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1} &= \begin{bmatrix} I & -A_{11}^{-1}A_{12} \\ 0 & I \end{bmatrix} \begin{bmatrix} A_{11}^{-1} & 0 \\ 0 & (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ -A_{21}A_{11}^{-1} & I \end{bmatrix} \\ &= \begin{bmatrix} A_{11}^{-1} + A_{11}^{-1}A_{12}(A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}(A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1} \\ -(A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1}A_{21}A_{11}^{-1} & (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1} \end{bmatrix} \end{aligned}$$

$$\begin{bmatrix} I & -A_{12}A_{22}^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} I & 0 \\ -A_{22}^{-1}A_{21} & I \end{bmatrix} = \begin{bmatrix} A_{11} - A_{12}A_{22}^{-1}A_{21} & 0 \\ 0 & A_{22} \end{bmatrix}$$

$$\begin{aligned} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1} &= \begin{bmatrix} I & 0 \\ -A_{22}^{-1}A_{21} & I \end{bmatrix} \begin{bmatrix} (A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1} & 0 \\ 0 & A_{22}^{-1} \end{bmatrix} \begin{bmatrix} I & -A_{12}A_{22}^{-1} \\ 0 & I \end{bmatrix} \\ &= \begin{bmatrix} (A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1} & -(A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1}A_{12}A_{22}^{-1} \\ -A_{22}^{-1}A_{21}(A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1} & A_{22}^{-1} + A_{22}^{-1}A_{21}(A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1}A_{12}A_{22}^{-1} \end{bmatrix} \end{aligned}$$

$$\begin{aligned} (A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1} &= A_{11}^{-1} + A_{11}^{-1}A_{12}(A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1}A_{21}A_{11}^{-1} \\ (A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1}A_{12}A_{22}^{-1} &= A_{11}^{-1}A_{12}(A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1} \\ (A_{11} + A_{12}A_{22}^{-1}A_{21})^{-1} &= A_{11}^{-1} - A_{11}^{-1}A_{12}(A_{22} + A_{21}A_{11}^{-1}A_{12})^{-1}A_{21}A_{11}^{-1} \\ (A_{11} + A_{12}A_{22}^{-1}A_{21})^{-1}A_{12}A_{22}^{-1} &= A_{11}^{-1}A_{12}(A_{22} + A_{21}A_{11}^{-1}A_{12})^{-1} \end{aligned}$$

Matrix Inversion Lemma

$$\begin{aligned}(A - BC^{-1}D)^{-1} &= A^{-1} + A^{-1}B(C - DA^{-1}B)^{-1}DA^{-1} \\ (A - BC^{-1}D)^{-1}BC^{-1} &= A^{-1}B(C - DA^{-1}B)^{-1} \\ (A + BC^{-1}D)^{-1} &= A^{-1} - A^{-1}B(C + DA^{-1}B)^{-1}DA^{-1} \\ (A + BC^{-1}D)^{-1}BC^{-1} &= A^{-1}B(C + DA^{-1}B)^{-1}\end{aligned}$$