Multivariate Gaussian

- A random vector $\mathbf{x} = [X_1 \ \ldots \ X_n]^T$ is said to be **multivariate Gaussian** if every linear combination of the components of X is a Gaussian random variable.
  - That is, for any a_i, $\sum_{i=1}^n a_i X_i$ is a Gaussian random variable.
  - We also say $X_1, \ldots, X_n$ are **jointly Gaussian**.
- **Multivariate Gaussian density function**:

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x}-\mu)^T \Sigma^{-1}(\mathbf{x}-\mu)\right) \tag{1}$$

$\mathbf{x} \sim \mathcal{N}(\mu, \Sigma)$, where μ is the mean vector and Σ is the covariance matrix.

$$\mu = \mathbb{E}(\mathbf{x}) = \begin{bmatrix} \mathbb{E}(X_1) \\ \vdots \\ \mathbb{E}(X_n) \end{bmatrix} \qquad \Sigma = \mathbf{cov}(\mathbf{x}) = \mathbb{E}\left(\left((\mathbf{x}-\mu)(\mathbf{x}-\mu)^T\right)\right)$$

Conditional Density of Multivariate Gaussian

Theorem: If $\mathbf{x} \in \mathbb{R}^r$ and $\mathbf{y} \in \mathbb{R}^m$ are jointly Gaussian with $n = r + m$, mean vector $[\mathbb{E}(\mathbf{x})^T \ \mathbb{E}(\mathbf{y})^T]^T$, and covariance matrix

$$\Sigma = \begin{bmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{bmatrix},$$

then the conditional probability density function $p(\mathbf{x}|\mathbf{y})$ is also a Gaussian random vector with mean $\mathbb{E}(\mathbf{x}|\mathbf{y})$ and covariance matrix $\Sigma_{x|y}$, where

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

Proof:

$$\begin{aligned} p(\mathbf{x}|\mathbf{y}) &= \frac{p(\mathbf{x},\mathbf{y})}{p(\mathbf{y})} && (1) \\ &= \frac{\frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}\begin{bmatrix} \mathbf{x} - \mathbb{E}(\mathbf{x}) \\ \mathbf{y} - \mathbb{E}(\mathbf{y}) \end{bmatrix}^T \Sigma^{-1} \begin{bmatrix} \mathbf{x} - \mathbb{E}(\mathbf{x}) \\ \mathbf{y} - \mathbb{E}(\mathbf{y}) \end{bmatrix}\right)}{\frac{1}{(2\pi)^{m/2}|\Sigma_{yy}|^{1/2}} \exp\left(-\frac{1}{2}[\mathbf{y} - \mathbb{E}(\mathbf{y})]^T \Sigma_{yy}^{-1} [\mathbf{y} - \mathbb{E}(\mathbf{y})]\right)} && (2) \end{aligned}$$

Proof Continued

$$
\begin{aligned}
p(\mathbf{x}|\mathbf{y}) &= \frac{\frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}\begin{bmatrix} \mathbf{x}-\mathbb{E}(\mathbf{x}) \\ \mathbf{y}-\mathbb{E}(\mathbf{y}) \end{bmatrix}^T \Sigma^{-1} \begin{bmatrix} \mathbf{x}-\mathbb{E}(\mathbf{x}) \\ \mathbf{y}-\mathbb{E}(\mathbf{y}) \end{bmatrix}\right)}{\frac{1}{(2\pi)^{m/2}|\Sigma_{yy}|^{1/2}} \exp\left(-\frac{1}{2}\left[\mathbf{y}-\mathbb{E}(\mathbf{y})\right]^T \Sigma_{yy}^{-1}\left[\mathbf{y}-\mathbb{E}(\mathbf{y})\right]\right)} && (1) \\
&= \frac{1}{(2\pi)^{r/2}\left(|\Sigma|/|\Sigma_{yy}|\right)^{1/2}} \exp\left(-\frac{1}{2}A\right), && (2)
\end{aligned}
$$

where

$$
A = \begin{bmatrix} \mathbf{x}-\mathbb{E}(\mathbf{x}) \\ \mathbf{y}-\mathbb{E}(\mathbf{y}) \end{bmatrix}^T \Sigma^{-1} \begin{bmatrix} \mathbf{x}-\mathbb{E}(\mathbf{x}) \\ \mathbf{y}-\mathbb{E}(\mathbf{y}) \end{bmatrix} - \left[\mathbf{y}-\mathbb{E}(\mathbf{y})\right]^T \Sigma_{yy}^{-1}\left[\mathbf{y}-\mathbb{E}(\mathbf{y})\right]. \quad (3)
$$

We now need to compute two terms:

- $|\Sigma|/|\Sigma_{yy}| = \det(\Sigma)/\det(\Sigma_{yy})$
- A

Proof Continued (Determinant)

We can easily verify that

$$\begin{bmatrix} I & -\Sigma_{xy}\Sigma_{yy}^{-1} \\ 0 & I \end{bmatrix} \underbrace{\begin{bmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{bmatrix}}_{\Sigma} \begin{bmatrix} I & 0 \\ -\Sigma_{yy}^{-1}\Sigma_{yx} & I \end{bmatrix} = \begin{bmatrix} \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx} & 0 \\ 0 & \Sigma_{yy} \end{bmatrix}$$

Hence,

$$\det(\Sigma) = \det(\Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx})\det(\Sigma_{yy}) \tag{1}$$

$$\frac{\det(\Sigma)}{\det(\Sigma_{yy})} = \det(\Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx}). \tag{2}$$

So we have

$$p(\mathbf{x}|\mathbf{y}) = \frac{1}{(2\pi)^{r/2}\left(|\Sigma|/|\Sigma_{yy}|\right)^{1/2}} \exp\left(-\frac{1}{2}A\right) \tag{3}$$

$$= \frac{1}{(2\pi)^{r/2}\left(\det(\Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx})\right)^{1/2}} \exp\left(-\frac{1}{2}A\right). \tag{4}$$

Proof Continued (A)

If $XYZ = W$ and matrices are invertible,

- By inverting both sides, we get $Z^{-1}Y^{-1}X^{-1} = W^{-1}$.
- Hence, $Y^{-1} = ZW^{-1}X$.

Since we have (where $S_{yy} = \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx}$)

$$\begin{bmatrix} I & -\Sigma_{xy}\Sigma_{yy}^{-1} \\ 0 & I \end{bmatrix} \underbrace{\begin{bmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{bmatrix}}_{\Sigma} \begin{bmatrix} I & 0 \\ -\Sigma_{yy}^{-1}\Sigma_{yx} & I \end{bmatrix} = \begin{bmatrix} S_{yy} & 0 \\ 0 & \Sigma_{yy} \end{bmatrix}, \quad (1)$$

$$\Sigma^{-1} = \begin{bmatrix} I & 0 \\ -\Sigma_{yy}^{-1}\Sigma_{yx} & I \end{bmatrix} \begin{bmatrix} S_{yy}^{-1} & 0 \\ 0 & \Sigma_{yy}^{-1} \end{bmatrix} \begin{bmatrix} I & -\Sigma_{xy}\Sigma_{yy}^{-1} \\ 0 & I \end{bmatrix}. \quad (2)$$

Now let $\mathbf{x}' = \mathbf{x} - \mathbb{E}(\mathbf{x})$ and $\mathbf{y}' = \mathbf{y} - \mathbb{E}(\mathbf{y})$. Then

$$\begin{aligned} A &= \begin{bmatrix} \mathbf{x}' \\ \mathbf{y}' \end{bmatrix}^T \Sigma^{-1} \begin{bmatrix} \mathbf{x}' \\ \mathbf{y}' \end{bmatrix} - \mathbf{y}'^T \Sigma_{yy}^{-1} \mathbf{y}'^T \quad (3) \\ &= \begin{bmatrix} \mathbf{x}' \\ \mathbf{y}' \end{bmatrix}^T \begin{bmatrix} I & 0 \\ -\Sigma_{yy}^{-1}\Sigma_{yx} & I \end{bmatrix} \begin{bmatrix} S_{yy}^{-1} & 0 \\ 0 & \Sigma_{yy}^{-1} \end{bmatrix} \begin{bmatrix} I & -\Sigma_{xy}\Sigma_{yy}^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} \mathbf{x}' \\ \mathbf{y}' \end{bmatrix} \\ &- \mathbf{y}'^T \Sigma_{yy}^{-1} \mathbf{y}'^T \quad (4) \end{aligned}$$

Proof Continued (A)

$$
\begin{aligned}
A &= \begin{bmatrix} \mathbf{x}' \\ \mathbf{y}' \end{bmatrix}^T \begin{bmatrix} I & 0 \\ -\Sigma_{yy}^{-1}\Sigma_{yx} & I \end{bmatrix} \begin{bmatrix} S_{yy}^{-1} & 0 \\ 0 & \Sigma_{yy}^{-1} \end{bmatrix} \begin{bmatrix} I & -\Sigma_{xy}\Sigma_{yy}^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} \mathbf{x}' \\ \mathbf{y}' \end{bmatrix} \\
&- \mathbf{y}'^T \Sigma_{yy}^{-1} \mathbf{y}'^T \\
&= \begin{bmatrix} \mathbf{x}' - \Sigma_{xy}\Sigma_{yy}^{-1}\mathbf{y}' \\ \mathbf{y}' \end{bmatrix}^T \begin{bmatrix} S_{yy}^{-1} & 0 \\ 0 & \Sigma_{yy}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{x}' - \Sigma_{xy}\Sigma_{yy}^{-1}\mathbf{y}' \\ \mathbf{y}' \end{bmatrix} - \mathbf{y}'^T \Sigma_{yy}^{-1} \mathbf{y}'^T \\
&= \left(\mathbf{x}' - \Sigma_{xy}\Sigma_{yy}^{-1}\mathbf{y}'\right)^T S_{yy}^{-1} \left(\mathbf{x}' - \Sigma_{xy}\Sigma_{yy}^{-1}\mathbf{y}'\right) \\
&= \left(\mathbf{x} - \left(\mathbb{E}(\mathbf{x}) + \Sigma_{xy}\Sigma_{yy}^{-1}(\mathbf{y} - \mathbb{E}(\mathbf{y}))\right)\right)^T S_{yy}^{-1} \left(\mathbf{x} - \left(\mathbb{E}(\mathbf{x}) + \Sigma_{xy}\Sigma_{yy}^{-1}(\mathbf{y} - \mathbb{E}(\mathbf{y}))\right)\right)
\end{aligned}
$$

where $S_{yy} = \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx}$.
Hence,

$$
p(\mathbf{x}|\mathbf{y}) = \frac{1}{(2\pi)^{r/2}|\Sigma_{x|y}|^{1/2}} \exp\left(-\frac{1}{2} (\mathbf{x} - \mathbb{E}(\mathbf{x}|\mathbf{y}))^T \Sigma_{x|y}^{-1} (\mathbf{x} - \mathbb{E}(\mathbf{x}|\mathbf{y})) \right), \quad (1)
$$

i.e., $\mathbf{x}|\mathbf{y} \sim \mathcal{N}(\mathbb{E}(\mathbf{x}|\mathbf{y}), \Sigma_{x|y})$, where

$$
\begin{aligned}
\mathbb{E}(\mathbf{x}|\mathbf{y}) &= \mathbb{E}(\mathbf{x}) + \Sigma_{xy}\Sigma_{yy}^{-1}(\mathbf{y} - \mathbb{E}(\mathbf{y})) && (2) \\
\Sigma_{x|y} &= \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx}. && (3)
\end{aligned}
$$