Course Notes	**Fall 2017**
Statistical Pattern Recognition (COMP136)	**Roni Khardon**
Department of Computer Science, Tufts University	

- **Partitioned Gaussians:** We divide a MVN variable $\mathbf{x}$ into two parts and calculate the conditional and marginal distributions on one part.

  $\mathbf{x} \sim \mathcal{N}(\mu, \Sigma)$ where we denote $\Lambda = \Sigma^{-1}$.

  The variables and parameters are partitioned as follows:

  $\mathbf{x} = \begin{pmatrix} \mathbf{x}_a \\ \mathbf{x}_b \end{pmatrix}$, $\mu = \begin{pmatrix} \mu_a \\ \mu_b \end{pmatrix}$, $\Sigma = \begin{pmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_{bb} \end{pmatrix}$, $\Lambda = \begin{pmatrix} \Lambda_{aa} & \Lambda_{ab} \\ \Lambda_{ba} & \Lambda_{bb} \end{pmatrix}$.

- The conditional:

$$\begin{aligned}
\mathbf{x}_a|(\mathbf{x}_b = x_b) &\sim \mathcal{N}(\mu_{a|b}, \Sigma_{a|b}) \\
\mu_{a|b} &= \mu_a - \Lambda_{aa}^{-1}\Lambda_{ab}(x_b - \mu_b) &&= \mu_a + \Sigma_{ab}\Sigma_{bb}^{-1}(x_b - \mu_b) \\
\Sigma_{a|b} &= \Lambda_{aa}^{-1} &&= \Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba}
\end{aligned}$$

- The marginal:

$$\begin{aligned}
\mathbf{x}_a &\sim \mathcal{N}(\mu_a^*, \Sigma_a^*) \\
\mu_a^* &= \mu_a \\
\Sigma_a^* &= [\Lambda_{aa} - \Lambda_{ab}\Lambda_{bb}^{-1}\Lambda_{ba}]^{-1} &&= \Sigma_{aa}
\end{aligned}$$

- **Linear combination of Gaussians:**
  $\mathbf{x} \sim \mathcal{N}(\mu, \Lambda^{-1})$ and $\mathbf{y} = A\mathbf{x} + b$ implies $\mathbf{y} \sim \mathcal{N}(A\mu + b, A\Lambda^{-1}A^T)$

- **Bayes theorem for linearly dependent Gaussians:** we have two variables distributed as $\mathbf{x} \sim \mathcal{N}(\mu, \Lambda^{-1})$ and $\mathbf{y}|\mathbf{x} \sim \mathcal{N}(A\mathbf{x} + b, L^{-1})$

- The joint distribution for $\mathbf{z} = \begin{pmatrix} x \\ y \end{pmatrix}$:

$$\begin{aligned}
\mathbf{z} &\sim \mathcal{N}(\mu_{\mathbf{z}}, \Lambda_{\mathbf{z}}^{-1}) \\
\mu_{\mathbf{z}} &= \begin{pmatrix} \mu \\ A\mu + b \end{pmatrix} \\
\Lambda_{\mathbf{z}}^{-1} &= \begin{pmatrix} \Lambda^{-1} & \Lambda^{-1}A^T \\ A\Lambda^{-1} & L^{-1} + A\Lambda^{-1}A^T \end{pmatrix}
\end{aligned}$$

- The conditional:

$$\begin{aligned}
\mathbf{x}|(\mathbf{y} = y) &\sim \mathcal{N}(\mu^*, \Sigma^*) \\
\Sigma^* &= (\Lambda + A^T L A)^{-1} \\
\mu^* &= \Sigma^*[A^T L(y - b) + \Lambda\mu]
\end{aligned}$$

- The marginal: $\mathbf{y} \sim \mathcal{N}(A\mu + b, L^{-1} + A\Lambda^{-1}A^T)$