Covariance Projection

November 9, 2014

Suppose there exists a random variable $x \sim \mathcal{N}(\mu_x, \Sigma_x)$ and we want to know the mean and covariance of $y = f(x)$. For the linear case where $y = Ax + b$,

$$\begin{aligned} \mathbb{E}[y] &= \mathbb{E}[Ax + b] \\ \mu_y &= A\mathbb{E}[x] + b \\ \mu_y &= A\mu_x + b \end{aligned}$$

$$\begin{aligned} \Sigma_{yy} &= \mathbb{E}[(y - \mathbb{E}[y])(y - \mathbb{E}[y])^T] \\ &= \mathbb{E}[(Ax + b - A\mu_x - b)(Ax + b - A\mu_x - b)^T] \\ &= \mathbb{E}[(Ax - A\mu_x)(Ax - A\mu_x)^T] \\ &= \mathbb{E}[A(x - \mu_x)(x - \mu_x)^T A^T] \\ &= A\mathbb{E}[(x - \mu_x)(x - \mu_x)^T]A^T \\ &= A\Sigma_x A^T \end{aligned}$$

$$\begin{aligned} \Sigma_{ab} &= \mathbb{E}[(a - \mathbb{E}[a])(b - \mathbb{E}[b])^T] \\ &= \mathbb{E}[(Ax + b - A\mu_x - b)(Bx + c - B\mu_x - c)^T] \\ &= \mathbb{E}[(Ax - A\mu_x)(Bx - B\mu_x)^T] \\ &= \mathbb{E}[A(x - \mu_x)(x - \mu_x)^T B^T] \\ &= A\mathbb{E}[(x - \mu_x)(x - \mu_x)^T]B^T \\ &= A\Sigma_x B^T \end{aligned}$$

To compute the relationship between y and any other variable z, given the covariance between x and z:

$$\begin{aligned} \Sigma_{yz} &= \mathbb{E}[(y - \mathbb{E}[y])(z - \mathbb{E}[z])^T] \\ &= \mathbb{E}[(Ax + b - A\mu_x - b)(z - \mu_z)^T] \\ &= \mathbb{E}[(Ax - A\mu_x)(z - \mu_z)^T] \\ &= \mathbb{E}[A(x - \mu_x)(z - \mu_z)^T] \\ &= A\mathbb{E}[(x - \mu_x)(z - \mu_z)^T] \\ &= A\Sigma_{xz} \end{aligned}$$

Similarly,

$$\begin{aligned}
\Sigma_{zy} &= \mathbb{E}[(z - \mathbb{E}[z])(y - \mathbb{E}[y])^T] \\
&= \mathbb{E}[(z - \mu_z)(Ax + b - A\mu_x - b)^T] \\
&= \mathbb{E}[(z - \mu_z)(Ax - A\mu_x)^T] \\
&= \mathbb{E}[(z - \mu_z)(x - \mu_x)^T A^T] \\
&= \mathbb{E}[(z - \mu_z)(x - \mu_x)^T] A^T \\
&= \Sigma_{zx} A^T
\end{aligned}$$

Now, assuming $y = Ax + b + n$ where $x \sim \mathcal{N}(\mu_x, \Sigma_x)$ and $n \sim \mathcal{N}(0, \Sigma_n)$, then the mean and covariance is given as

$$\begin{aligned}
\mathbb{E}[y] &= \mathbb{E}[Ax + b + n] \\
\mu_y &= A\mathbb{E}[x] + b + \mathbb{E}[n] \\
\mu_y &= A\mu_x + b + 0 \\
\mu_y &= A\mu_x + b
\end{aligned}$$

$$\begin{aligned}
\Sigma_{yy} &= \mathbb{E}[(y - \mathbb{E}[y])(y - \mathbb{E}[y])^T] \\
&= \mathbb{E}[(Ax + b + n - A\mu_x - b)(Ax + b + n - A\mu_x - b)^T] \\
&= \mathbb{E}[(Ax - A\mu_x + n)(Ax - A\mu_x + n)^T] \\
&= \mathbb{E}[A(x - \mu_x)(x - \mu_x)^T A^T + n(x - \mu_x)^T A^T + A(x - \mu_x)n + nn^T] \\
&= A\mathbb{E}[(x - \mu_x)(x - \mu_x)^T]A^T + \mathbb{E}[n(x - \mu_x)^T A^T] + \mathbb{E}[A(x - \mu_x)n] + \mathbb{E}[nn^T] \\
&= A\Sigma_x A^T + \Sigma_n
\end{aligned}$$

Since $\mathbb{E}[n(x - \mu_x)^T A^T] = \mathbb{E}[n]\mathbb{E}[(x - \mu_x)^T A^T] = 0$

$$\begin{aligned}
\Sigma_{ab} &= \mathbb{E}[(a - \mathbb{E}[a])(b - \mathbb{E}[b])^T] \\
&= \mathbb{E}[(Ax + b + n_1 - A\mu_x - b)(Bx + c + n_2 - B\mu_x - c)^T] \\
&= \mathbb{E}[(Ax - A\mu_x + n_1)(Bx - B\mu_x + n_2)^T] \\
&= \mathbb{E}[A(x - \mu_x)(x - \mu_x)^T B^T + n_1(x - \mu_x)^T B^T + A(x - \mu_x)n_2 + n_1 n_2^T] \\
&= A\mathbb{E}[(x - \mu_x)(x - \mu_x)^T]B^T + 0 + 0 + 0 \\
&= A\Sigma_x B^T
\end{aligned}$$