2.2.4 Statistically Independent, Uncorrelated

In the case of Gaussian PDFs, statistically independent variables are also uncorrelated (true in general) and uncorrelated variables are also statistically independent (not true for all types of PDFs). We can see this fairly easily by looking at (2.53). If we assume statistical independence, $p(\mathbf{x}, \mathbf{y}) = p(\mathbf{x})p(\mathbf{y})$ and so $p(\mathbf{x}|\mathbf{y}) = p(\mathbf{x}) = \mathcal{N}(\boldsymbol{\mu}_x, \boldsymbol{\Sigma}_{xx})$. Looking at (2.53b), this implies

$$\boldsymbol{\Sigma}_{xy}\boldsymbol{\Sigma}_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y) = \mathbf{0}, \tag{2.54a}$$

$$\boldsymbol{\Sigma}_{xy}\boldsymbol{\Sigma}_{yy}^{-1}\boldsymbol{\Sigma}_{yx} = \mathbf{0}, \tag{2.54b}$$

which further implies that $\boldsymbol{\Sigma}_{xy} = \mathbf{0}$. Since

$$\boldsymbol{\Sigma}_{xy} = E\left[(\mathbf{x} - \boldsymbol{\mu}_x)(\mathbf{y} - \boldsymbol{\mu}_y)^T\right] = E\left[\mathbf{x}\mathbf{y}^T\right] - E\left[\mathbf{x}\right]E\left[\mathbf{y}\right]^T, \tag{2.55}$$

we have the uncorrelated condition:

$$E\left[\mathbf{x}\mathbf{y}^T\right] = E\left[\mathbf{x}\right]E\left[\mathbf{y}\right]^T. \tag{2.56}$$

We can also work through the logic in the other direction by first assuming the variables are uncorrelated, which leads to $\boldsymbol{\Sigma}_{xy} = \mathbf{0}$, and finally to statistical independence. Since these conditions are equivalent, we will often use *statistically independent* and *uncorrelated* interchangeably in the context of Gaussian PDFs.

2.2.5 Linear Change of Variables

Suppose that we have a Gaussian random variable,

$$\mathbf{x} \in \mathbb{R}^N \sim \mathcal{N}(\boldsymbol{\mu}_x, \boldsymbol{\Sigma}_{xx}),$$

and that we have a second random variable, $\mathbf{y} \in \mathbb{R}^M$, related to $\mathbf{x}$ through the linear map,

$$\mathbf{y} = \mathbf{G}\mathbf{x}, \tag{2.57}$$

where we assume that $\mathbf{G} \in \mathbb{R}^{M\times N}$ is a constant matrix. We would like to know what the statistical properties of $\mathbf{y}$ are. One way to do this is to simply apply the expectation operator directly:

$$\boldsymbol{\mu}_y = E[\mathbf{y}] = E[\mathbf{G}\mathbf{x}] = \mathbf{G}\,E[\mathbf{x}] = \mathbf{G}\boldsymbol{\mu}_x, \tag{2.58a}$$

$$\begin{aligned}\boldsymbol{\Sigma}_{yy} &= E[(\mathbf{y} - \boldsymbol{\mu}_y)(\mathbf{y} - \boldsymbol{\mu}_y)^T] \\ &= \mathbf{G}\,E[(\mathbf{x} - \boldsymbol{\mu}_x)(\mathbf{x} - \boldsymbol{\mu}_x)^T]\,\mathbf{G}^T = \mathbf{G}\boldsymbol{\Sigma}_{xx}\mathbf{G}^T,\end{aligned} \tag{2.58b}$$

so that we have $\mathbf{y} \sim \mathcal{N}(\boldsymbol{\mu}_y, \boldsymbol{\Sigma}_{yy}) = \mathcal{N}(\mathbf{G}\boldsymbol{\mu}_x, \mathbf{G}\boldsymbol{\Sigma}_{xx}\mathbf{G}^T)$.

Another way to look at this is a change of variables. We assume that the linear map is *injective*, meaning two $\mathbf{x}$ values cannot map to a single $\mathbf{y}$ value; in fact, let us simplify the injective condition by assuming a