Multivariate Normal - Cheat Sheet

Matthew Pearce

June 22, 2015

1 Covariance parameterisation

1.1 Partitioned vector

$$\mathbf{z} = \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix}, \begin{bmatrix} \mathbf{\Sigma}_{aa} & \mathbf{\Sigma}_{ab} \\ \mathbf{\Sigma}_{ab}^T & \mathbf{\Sigma}_{bb} \end{bmatrix} \right)$$

1.2 Marginals

$$\mathbf{x} \sim \mathcal{N}\left(\mathbf{a}, \mathbf{\Sigma}_{aa}\right)$$

1.3 Conditionals

$$\mathbf{x}|\mathbf{y} \sim \mathcal{N}\left(\mathbf{a} + \mathbf{\Sigma}_{ab}\mathbf{\Sigma}_{bb}^{-1}(\mathbf{y} - \mathbf{b}), \mathbf{\Sigma}_{aa} - \mathbf{\Sigma}_{ab}\mathbf{\Sigma}_{bb}^{-1}\mathbf{\Sigma}_{ab}^T\right)$$

1.4 x as mean for many IID observations

Suppose that $\mathbf{x}$ serves as the mean for N IID observations $\mathbf{y}_1, ..., \mathbf{y}_N$:

$$\begin{aligned} &\mathbf{x} \sim \mathcal{N}(\mathbf{a}, \Theta) \\ &\mathbf{y}_i|\mathbf{x} \sim \mathcal{N}(\mathbf{x}, \Sigma)\ ,\ i = 1, ..., N \end{aligned}$$

Then the posterior distribution on $\mathbf{x}$ is normal with:

$$\mathbf{x}|\mathbf{y}_1, ..., \mathbf{y}_N \sim \mathcal{N}((\Theta^{-1} + N\Sigma^{-1})^{-1}(\Theta^{-1}\mathbf{a} + N\Sigma^{-1}\bar{\mathbf{y}}), (\Theta^{-1} + N\Sigma^{-1})^{-1})$$

2 Precision parameterisation

2.1 Partitioned vector

$$\mathbf{z} = \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix}, \begin{bmatrix} \mathbf{\Lambda}_{aa} & \mathbf{\Lambda}_{ab} \\ \mathbf{\Lambda}_{ab}^T & \mathbf{\Lambda}_{bb} \end{bmatrix}^{-1} \right)$$

Using the block inverse we can find the covariance matrix:

$$\begin{bmatrix} \mathbf{\Lambda}_{aa} & \mathbf{\Lambda}_{ab} \\ \mathbf{\Lambda}_{ab}^T & \mathbf{\Lambda}_{bb} \end{bmatrix}^{-1} = \begin{bmatrix} (\mathbf{\Lambda}_{aa} - \mathbf{\Lambda}_{ab}\mathbf{\Lambda}_{bb}^{-1}\mathbf{\Lambda}_{ab}^T)^{-1} & -\mathbf{\Lambda}_{aa}^{-1}\mathbf{\Lambda}_{ab}(\mathbf{\Lambda}_{bb} - \mathbf{\Lambda}_{ab}^T\mathbf{\Lambda}_{aa}^{-1}\mathbf{\Lambda}_{ab})^{-1} \\ -\mathbf{\Lambda}_{bb}^{-1}\mathbf{\Lambda}_{ab}^T(\mathbf{\Lambda}_{aa} - \mathbf{\Lambda}_{ab}\mathbf{\Lambda}_{bb}^{-1}\mathbf{\Lambda}_{ab}^T)^{-1} & (\mathbf{\Lambda}_{bb} - \mathbf{\Lambda}_{ab}^T\mathbf{\Lambda}_{aa}^{-1}\mathbf{\Lambda}_{ab})^{-1} \end{bmatrix}$$
$$= \begin{bmatrix} \mathbf{\Sigma}_{aa} & \mathbf{\Sigma}_{ab} \\ \mathbf{\Sigma}_{ab}^T & \mathbf{\Sigma}_{bb} \end{bmatrix}$$

2.2 Marginals

$$\mathbf{x} \sim \mathcal{N}\left(\mathbf{a}, \quad \left(\mathbf{\Lambda}_{aa} - \mathbf{\Lambda}_{ab}\mathbf{\Lambda}_{bb}^{-1}\mathbf{\Lambda}_{ab}^T \right)^{-1} \right)$$

2.3 Conditionals

$$\mathbf{x}|\mathbf{y} \sim \mathcal{N}\left(\mathbf{a} + \mathbf{\Lambda}_{aa}^{-1}\mathbf{\Lambda}_{ab}(\mathbf{y} - \mathbf{b}), \quad \mathbf{\Lambda}_{aa}^{-1} \right)$$

2.4 Reverse conditionals

Where $\mathbf{A}$ is a matrix of constants (often provided by a given covariance matrix), and we are given two random vectors along with a prior for one and a conditional/likelihood for the other. See Bishop PRML p.93:

$$\mathbf{y} \sim \mathcal{N}(\mathbf{b}, \quad \mathbf{\Lambda}_{bb}^{-1})$$
$$\mathbf{x}|\mathbf{y} \sim \mathcal{N}\left(\mathbf{A}\mathbf{y} + \mathbf{a}, \quad \mathbf{\Lambda}_{aa}^{-1} \right)$$

Then we can obtain the reverse conditional and the other marginal:

$$\mathbf{x} \sim \mathcal{N}(\mathbf{A}\mathbf{b} + \mathbf{a}, \quad \mathbf{\Lambda}_{aa}^{-1} + \mathbf{A}\mathbf{\Lambda}_{bb}^{-1}\mathbf{A}^T)$$
$$\mathbf{y}|\mathbf{x} \sim \mathcal{N}\left(\left(\mathbf{\Lambda}_{bb} + \mathbf{A}^T\mathbf{\Lambda}_{aa}\mathbf{A} \right)^{-1} \left(\mathbf{A}^T\mathbf{\Lambda}_{aa}(\mathbf{x} - \mathbf{a}) + \mathbf{\Lambda}_{bb}\mathbf{b} \right), \quad \left(\mathbf{\Lambda}_{bb} + \mathbf{A}^T\mathbf{\Lambda}_{aa}\mathbf{A} \right)^{-1} \right)$$

3 Making Bibtex happy

Some foo by Other (2000).

References

Amy N. Other. Dummy. *Journal of foo*, 1(1):1–10, 2000.