

Multivariate Normal - Cheat Sheet

Matthew Pearce

December 9, 2015

1 Covariance parameterisation

1.1 Partitioned vector

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

1.2 Marginals

$$\mathbf{x} \sim \mathcal{N}(\mathbf{a}, \Sigma_{aa})$$

1.3 Conditionals

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

1.4 \mathbf{x} as mean for many IID observations

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

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

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

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

2 Precision parameterisation

2.1 Partitioned vector

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

Using the block inverse we can find the covariance matrix:

$$\begin{aligned} \begin{bmatrix} \Lambda_{aa} & \Lambda_{ab} \\ \Lambda_{ab}^T & \Lambda_{bb} \end{bmatrix}^{-1} &= \begin{bmatrix} (\Lambda_{aa} - \Lambda_{ab}\Lambda_{bb}^{-1}\Lambda_{ab}^T)^{-1} & -\Lambda_{aa}^{-1}\Lambda_{ab}(\Lambda_{bb} - \Lambda_{ab}^T\Lambda_{aa}^{-1}\Lambda_{ab})^{-1} \\ -\Lambda_{bb}^{-1}\Lambda_{ab}^T(\Lambda_{aa} - \Lambda_{ab}\Lambda_{bb}^{-1}\Lambda_{ab}^T)^{-1} & (\Lambda_{bb} - \Lambda_{ab}^T\Lambda_{aa}^{-1}\Lambda_{ab})^{-1} \end{bmatrix} \\ &= \begin{bmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ab}^T & \Sigma_{bb} \end{bmatrix} \end{aligned}$$

2.2 Marginals

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

2.3 Conditionals

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

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:

$$\begin{aligned} \mathbf{y} &\sim \mathcal{N}(\mathbf{b}, \Lambda_{bb}^{-1}) \\ \mathbf{x}|\mathbf{y} &\sim \mathcal{N}(\mathbf{A}\mathbf{y} + \mathbf{a}, \Lambda_{aa}^{-1}) \end{aligned}$$

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

$$\begin{aligned} \mathbf{x} &\sim \mathcal{N}(\mathbf{A}\mathbf{b} + \mathbf{a}, \Lambda_{aa}^{-1} + \mathbf{A}\Lambda_{bb}^{-1}\mathbf{A}^T) \\ \mathbf{y}|\mathbf{x} &\sim \mathcal{N} \left((\Lambda_{bb} + \mathbf{A}^T\Lambda_{aa}\mathbf{A})^{-1} (\mathbf{A}^T\Lambda_{aa}(\mathbf{x} - \mathbf{a}) + \Lambda_{bb}\mathbf{b}), (\Lambda_{bb} + \mathbf{A}^T\Lambda_{aa}\mathbf{A})^{-1} \right) \end{aligned}$$

3 Making Bibtex happy

Some foo by Other (2000).

References

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