086761 - Homework 1

Yuri Feldman, 309467801

November 6, 2017

1

1.1

$$f_x(x) = \frac{1}{\sqrt{|2\pi\Sigma_x|}} \cdot e^{-\frac{1}{2}(x-\mu_x)^T\Sigma_x^{-1}(x-\mu_x)} \tag{1.1}$$

1.2

μ_y, Σ_y can be calculated using expectation properties (and do not depend on y being Gaussian):

$$\mu_y = \mathbb{E}Ax + b = A\mathbb{E}x + b = \boxed{A\mu_x + b} \tag{1.2}$$

$$\Sigma_y = \mathbb{E}\left[(Ax + b - \mu_y)(Ax + b - \mu_y)^T\right] = \mathbb{E}\left[A(x-\mu_x)(x-\mu_x)^T A^T\right] = \tag{1.3}$$

$$A\mathbb{E}\left[(x-\mu_x)(x-\mu_x)^T\right]A^T = \boxed{A\Sigma_x A^T} \tag{1.4}$$

Next, we show that y is Gaussian. In the case where A is invertible this can be directly obtained from random vector transformation formula (1-to-1 transformation). Since $y = Ax + b$ the Jacobian matrix is A and since the mapping is one-to-one (A is invertible) according to the transformation theorem:

$$f_Y(y) = \frac{1}{|A|} f_X(A^{-1}(y-b)) \tag{1.5}$$

ignoring the constant normalization factors (since they don't depend on x and only guarantee the distributions' summing to 1) we calculate the exponent in the right hand side:

$$-\frac{1}{2}\left(A^{-1}(y-b) - \mu_x\right)^T \Sigma_x^{-1}\left(A^{-1}(y-b) - \mu_x\right) = \tag{1.6}$$

$$-\frac{1}{2}\left(y - b - A\mu_x\right)^T A^{-T}\Sigma_x^{-1}A^{-1}\left(y - b - A\mu_x\right) = \tag{1.7}$$

$$-\frac{1}{2}\left(y - (A\mu_x + b)\right)^T \left(A\Sigma_x A^T\right)^{-1}\left(y - (A\mu_x + b)\right), \tag{1.8}$$

which exactly gives a Gaussian distribution (in particular, with the mean and variance we calculated beforehand).

We now deal with the case where $A \in \mathbb{R}^{m\times n}$ is not invertible - either rectangular or square with zero determinant. Consider the SVD decomposition of A:

$$A = U\Sigma V^T \tag{1.9}$$

Here, U and V are square invertible matrices of respective sizes $m\times m$ and $n\times n$. Note that $V^T = V^*$ as A is real. Σ is a rectangular matrix the same size as A with the main diagonal containing the singular values (and zero elsewhere). We can now write:

$$y = U\Sigma V^T x + b \tag{1.10}$$

Here, $V^T x$ is Gaussian, since V is invertible. Without loss of generality denote

$$\Sigma = \begin{pmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{pmatrix} \qquad V^T x = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}, \tag{1.11}$$

where Σ_1 is (square) diagonal, with positive entries (and in particular, is invertible), and has the same number of columns as the length of vector v_1. We have then that

$$\Sigma V^T x = \begin{pmatrix} \Sigma_1 v_1 \\ 0 \end{pmatrix}, \tag{1.12}$$

where $\Sigma_1 v_1$ is Gaussian since v_1 is and Σ_1 is invertible, and hence $\Sigma V^T x = U^T(y-b)$ is (possibly degenerate) Gaussian (and hence, so is y).

2

2.1

$$p(x \mid z) = \frac{p(z \mid x)p(x)}{p(z)} = \frac{p(z \mid x)p(x)}{\int_x p(z \mid x)p(x) \cdot dx} \propto p(z \mid x) \cdot p(x) \tag{2.1}$$

2.2

We are interested in the MAP estimate:

$$x^*_{MAP} \doteq \arg\max_x p(x \mid z) = \arg\max_x p(z \mid x) \cdot p(x) = \arg\max_x \log p(z \mid x) + \log p(x) = \tag{2.2}$$

$$\arg\min_x ||z - Hx||^2_R + ||x - x_0||^2_{\Sigma_0} \tag{2.3}$$

Develop the latter:

$$||z - Hx||^2_R + ||x - x_0||^2_{\Sigma_0} = \left|\left| R^{-1/2}(Hx - z) \right|\right|^2 + \left|\left| \Sigma_0^{-1/2}(x - x_0) \right|\right|^2 = \tag{2.4}$$

$$\left|\left| \begin{pmatrix} R^{-1/2}(Hx - z) \\ \Sigma_0^{-1/2}(x - x_0) \end{pmatrix} \right|\right|^2 = \left|\left| \begin{pmatrix} R^{-1/2}H \\ \Sigma_0^{-1/2} \end{pmatrix} x - \begin{pmatrix} R^{-1/2}z \\ \Sigma_0^{-1/2}x_0 \end{pmatrix} \right|\right|^2, \tag{2.5}$$

from where x^*_{MAP} can be obtained as the least-squares solution using the pseudoinverse $A^\dagger b = (A^T A)^{-1}A^T b$, or equivalently proceed directly, noting that the function is convex and has a unique extremum which is the minimum. Develop the above into:

$$= x^T \left(H^T R^{-1} H + \Sigma_0^{-1}\right) x - 2\left(z^T R^{-1} H + x_0^T \Sigma_0^{-1}\right) x + \left(z^T R^{-1} z + x_0^T \Sigma_0^{-1} x_0\right) \tag{2.6}$$

find the zero of the gradient:

$$\nabla(\cdot) = 2\left(H^T R^{-1} H + \Sigma_0^{-1}\right) x^* - 2\left(\Sigma_0^{-1} x_0 + H^T R^{-1} z\right) = 0 \tag{2.7}$$

$$\implies \boxed{x^*_{MAP} = \left(H^T R^{-1} H + \Sigma_0^{-1}\right)^{-1} \left(\Sigma_0^{-1} x_0 + H^T R^{-1} z\right)} \tag{2.8}$$

The associated covariance is

$$\boxed{\Sigma = \left(H^T R^{-1} H + \Sigma_0^{-1}\right)^{-1}}, \tag{2.9}$$

as can be seen from the quadratic term in Eq.(2.6).

3 Code

$$\begin{pmatrix} \cos(\psi)\cos(\theta) & \cos(\psi)\sin(\phi)\sin(\theta) - \cos(\phi)\sin(\psi) & \sin(\phi)\sin(\psi) + \cos(\phi)\cos(\psi)\sin(\theta) \\ \cos(\theta)\sin(\psi) & \cos(\phi)\cos(\psi) + \sin(\phi)\sin(\psi)\sin(\theta) & \cos(\phi)\sin(\psi)\sin(\theta) - \cos(\psi)\sin(\phi) \\ -\sin(\theta) & \cos(\theta)\sin(\phi) & \cos(\phi)\cos(\theta) \end{pmatrix} \tag{3.1}$$

For the case $\theta = \frac{\pi}{2}$:

$$\begin{pmatrix} 0 & \cos(\psi)\sin(\phi) - \cos(\phi)\sin(\psi) & \sin(\phi)\sin(\psi) + \cos(\phi)\cos(\psi) \\ 0 & \cos(\phi)\cos(\psi) + \sin(\phi)\sin(\psi) & \cos(\phi)\sin(\psi) - \cos(\psi)\sin(\phi) \\ -1 & 0 & 0 \end{pmatrix} \tag{3.2}$$

$$= \begin{pmatrix} 0 & \sin(\phi - \psi) & \cos(\phi - \psi) \\ 0 & \cos(\phi - \psi) & -\sin(\phi - \psi) \\ -1 & 0 & 0 \end{pmatrix} \tag{3.3}$$

For the case $\theta = -\frac{\pi}{2}$:

$$\begin{pmatrix} 0 & -\cos(\psi)\sin(\phi) - \cos(\phi)\sin(\psi) & \sin(\phi)\sin(\psi) - \cos(\phi)\cos(\psi) \\ 0 & \cos(\phi)\cos(\psi) - \sin(\phi)\sin(\psi) & -\cos(\phi)\sin(\psi) - \cos(\psi)\sin(\phi) \\ 1 & 0 & 0 \end{pmatrix} \tag{3.4}$$

$$= \begin{pmatrix} 0 & -\sin(\phi + \psi) & -\cos(\phi + \psi) \\ 0 & \cos(\phi + \psi) & -\sin(\phi + \psi) \\ 1 & 0 & 0 \end{pmatrix} \tag{3.5}$$