Homework 2

Tyler Amos, 10 May 2019

1a

To prove that covariance matrices are symmetric, it is sufficient to show:

$$\Sigma_{ij} = \Sigma_{ji}, \forall i, j$$
$$\Sigma = \Sigma^T$$
$$E[(x - E(x))(x - E(x))^T] = E[(x - E(x))(x - E(x))^T]^T$$

Let's express $(x - E(x))$ as A:

$$E[AA^T] = E[AA^T]^T$$

Now because for any matrices x and y:

$$(xy)^T = y^T x^T$$

We can write the above expression as:

$$E[AA^T] = E[AA^T]$$

1b

We know from 1a that covariate matrices are symmetric. So

$$A = A^T$$

if A is a symmetric matrix. We start by defining an equality for the condition of positive semi-definite matrices.

Claim: For the covariance matrix M and the vector c:

$$c^T Mc > 0, \forall c \in R^n$$

Substitute M for the definition of a covariance matrix:

$$c^T E[(x - E(x))(x - E(x))^T]c > 0$$

Substitute $(x - E(x)) = b$ to simplify notation:

$$c^T E[b(b^T)]c > 0$$

Since b is symmetric:

$$c^T E[b^2]c > 0$$

b^2 with always be positive, so:

$$c^T E[b^2]c > 0, \forall c \in R^n$$

Thus symmetry guarantees a covariance matrix will be positive semi-definite.

1c

1d

You could approach this problem using a weighted sampling method. More concretely, if you specify the function to which you do have access as f, and the underlying function to which it is related by a linear map as g, then we can approximate g from f by iteratively re-weighting our samples from f, where the weight is proportional to the likelihood of the sample value if the distribution is gaussian.

1e

Recall that gaussians are closed under linear transformations so the mean and variance are simply linear transformations of the parameters of an untransformed multivariate gaussian.

2a

I assume H in the question refers to a constant (i.e., the letter) not capital η

First let us examine the mean of Z, μ_z:

$$\mu_z = Z(x(\hat{x}^-; \hat{\Sigma}^-))$$

By the fact that gaussians are closed under linear transformation the mean is thus:

$$= H\hat{x}^- + v$$

Similarly, the covariance can be derived as:

$$\Sigma_z = E[(\mu_z - E[\mu_z]) - (\mu_z - E[\mu_z])^T]$$

Where

$$\mu_z = H\hat{x}^- + v$$

2b

Claim:

$$p(x|z) = \frac{1}{p(z)} p(v)p(x)$$

By Bayes' Theorem:

$$p(x|z) = \frac{1}{p(z)} p(z|x)p(x)$$

where $\frac{1}{p(z)}$ is a constant scaling factor, so it suffices to show:

$$p(z|x)p(x) = p(x)p(v)$$

Note that:

$$p(z|x) = p(Hx + v)p(x)$$

This implies the additional values supplied by the term Hx will consistently vary , 1:1, with the term x in $p(x)$. This allow us to drop these terms as they provide no additional information beyond that provided by x. This implies:

$$p(z|x)p(x) = p(Hx + v)p(x) = p(v)p(x)$$

2c

In order to find j in:

$$p(x|z) \propto \exp(j)$$

First recall:

$$\mu_z = H\hat{x}^- + v$$
$$\Sigma_z = E[(\mu_z - E[\mu_z]) - (\mu_z - E[\mu_z])^T]$$

And the multivariate normal is:

$$\det(2\pi\Sigma)^{-\frac{1}{2}} \exp(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu))$$

Substituting $\mu = \mu_z, \Sigma^{-1} = \Sigma_z$ we get our expression for j in the exponentiation of:

$$\det(2\pi\Sigma)^{-\frac{1}{2}} \exp(-\frac{1}{2}(x - H\hat{x}^- + v)^T(E[(\mu_z - E[\mu_z]) - (\mu_z - E[\mu_z])^T])^{-1}(x - H\hat{x}^- + v))$$

2d

The expression derived above is cumbersome, so we simplify. First, we restate the result in terms of the original statement in 2c:

$$p(x|z) \propto \exp(-\frac{1}{2}(x - H\hat{x}^- + v)^T(E[(\mu_z - E[\mu_z]) - (\mu_z - E[\mu_z])^T])^{-1}(x - H\hat{x}^- + v))$$

Which we will try to simplify to:

$$p(x|z) \propto \exp(-\frac{1}{2}(x - \hat{x}^+)^T(E[(H\hat{x}^- + v - E[H\hat{x}^- + v]) - (H\hat{x}^- + v - E[H\hat{x}^- + v])^T])^{+-1}(x - \hat{x}^+))$$

Focusing just on the exponentiated term, we can rearrange this as:

$$-\frac{(x - H\hat{x}^- + v)^T(x - H\hat{x}^- + v)}{(2)E[(H\hat{x}^- + v - E[H\hat{x}^- + v]) - (H\hat{x}^- + v - E[H\hat{x}^- + v])^T])}$$

2e

Not completed.

3a

Not completed.

4a

The Jacobians can be expressed as:

$$F = \begin{bmatrix} 1 & -d_t sin(\theta_{t-1}) & 0 \\ 0 & 1 & dcos(\theta_{t-1}) \\ 0 & 0 & 1 \end{bmatrix} \tag{1}$$

$$H = \begin{bmatrix} 2x & 2y & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2}$$

4b

See attached files for code and plots. The final covariance matrix and mean vector summary for the original parametrization is (also available in final_covariance_and_mean.txt):

Final mean vector:

$$\begin{bmatrix} 2.82761774 & -4.34687978 & 1.47964579 \end{bmatrix} \tag{3}$$

Final covariance matrix:

$$\begin{bmatrix} 1.85849588e-02 & 1.21165287e-02 & 5.50086782e-14 \\ 1.21165287e-02 & 7.89942626e-03 & 3.58668628e-14 \\ 5.50086782e-14 & 3.58668628e-14 & 1.73668880e-08 \end{bmatrix} \tag{4}$$

Plots in the 'classic' directory are those which used the specified Q and R values from the assignment. Those in the 'custom' and qr_* directories are various parameterizations. For an example of a parameterization which suggests the estimator is overconfident, consider:

$$R = \begin{bmatrix} 2.0 & 0.0 & 0.0 \\ 0.0 & 2.0 & 0.0 \\ 0.0 & 0.0 & (2.0 \times \pi)/180 \end{bmatrix} \times \frac{1}{21} \tag{5}$$

$$Q = \begin{bmatrix} 1.0 & 0.0 \\ 0.0 & \pi/180 \end{bmatrix} \times \frac{1}{21} \tag{6}$$

The corresponding plots are provided in Figures 1, 2.

5a

I did not collaborate with anyone on this problem set.

5b

I spent approximately 13 hours on the problem set.

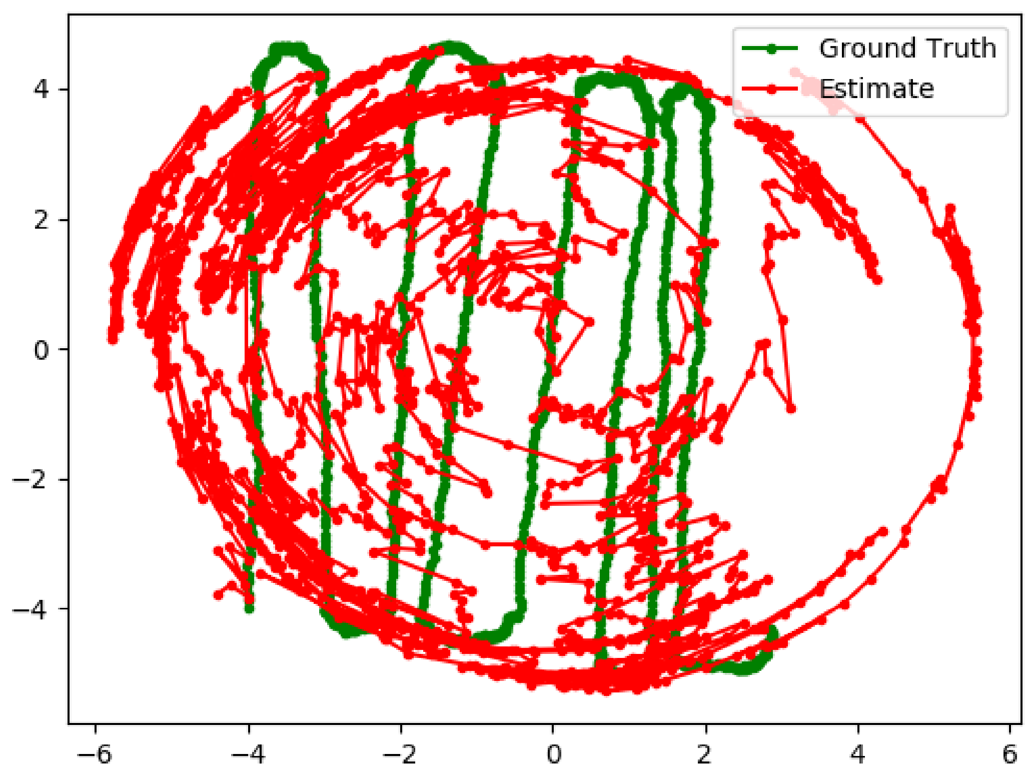

Figure 1: Trajectory Plot

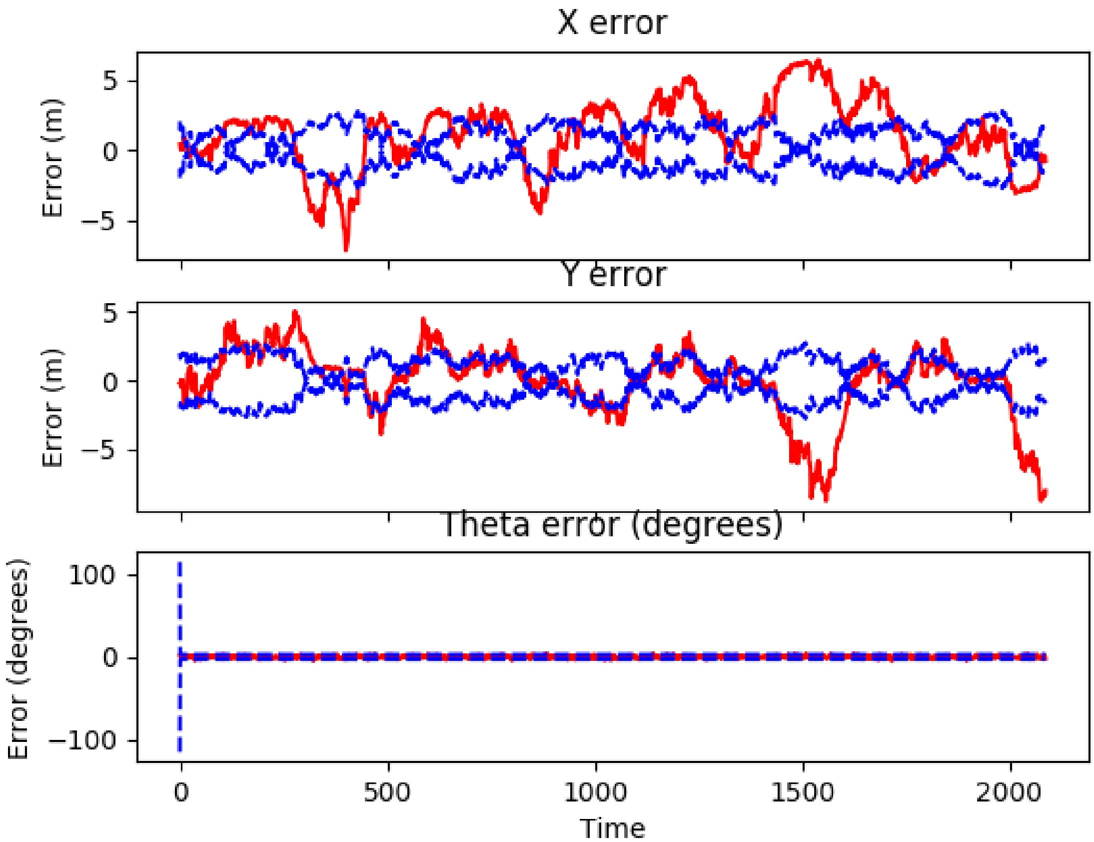

Figure 2: Error Plot