Problem Set 1

1. In the scalar example of section 1.6, show that the analysis error is uncorrelated with the observation increment, i.e.
$$< \epsilon^a(\epsilon^{\text{obs}} - \epsilon^b) >= 0.$$

2. In the simple scalar example, suppose that the observation and background errors are now biased, i.e.
$$< \epsilon^{\text{obs}} >= b^{\text{obs}}, \quad < \epsilon^b >= b^b.$$
Construct new variables that are unbiased:
$$\begin{aligned} \tilde{\mathrm{x}}^{\text{obs}} &= \mathrm{x}^{\text{obs}} - b^{\text{obs}} \\ \tilde{\mathrm{x}}^b &= \mathrm{x}^b - b^b. \end{aligned} \tag{1}$$
   (a) Form the analysis equation in terms of errors.
   (b) Show that $< \mathrm{x}^a - \mathrm{x}^t >= 0$.
   (c) Find W that minimizes $< (\epsilon^a)^2 >$.

3. a) Consider a Gaussian variable, x, with mean μ and variance σ^2. Let $\mathrm{y} = a\mathrm{x} + b$. What is the pd.f. of y?

   b) Show that the linear transformation of a normally distributed vector is also normally distributed. That is, show that for a given normally distributed vector $\mathbf{x}$, with mean $\mu_{\mathbf{x}}$ and covariance $\mathbf{R_x}$, the linear transformation
$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{b}$$
   produces a normally distributed vector $\mathbf{y}$ with mean $\mu_{\mathbf{y}} = \mathbf{A}\mu_{\mathbf{x}} + \mathbf{b}$ and covariance $\mathbf{R_y} = \mathbf{A}\mathbf{R_x}\mathbf{A}^T$.

4. Consider the following three distributions:

   1) Uniform:
$$p_{\mathrm{x}}(x) = \begin{cases} \frac{1}{2\sqrt{3}}, & \text{if } -\sqrt{3} \leq x \leq \sqrt{3} \\ 0, & \text{otherwise} \end{cases} \tag{2}$$
   2) Triangular:
$$p_{\mathrm{x}}(x) = \begin{cases} \frac{(\sqrt{6}+x)}{6}, & \text{if } -\sqrt{6} \leq x \leq 0 \\ \frac{(\sqrt{6}-x)}{6}, & \text{if } 0 < x \leq \sqrt{6} \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

3) Gaussian, where x is $\mathcal{N}(0,1)$:

$$p_{\rm x}(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2}\right\} \tag{4}$$

a) Plot these curves for x=-4.0:0.1:4.0 .

```
x=-4.0:.1:4.0;
plot(x,uniform(x),x,triang(x),x,gauss(0,1,x))
```

The script names should end in ".m". As an example, here is gauss.m

```
function y=gauss(mean,sig,x)
n=size(x,2);
y=zeros(1,n);
y=(1/sqrt(2*pi*sig)).*exp(-0.5.*((x-mean)/sig).^2);
```

Note that a dot before an operator indicates a matrix operation.

b) What is the mean and variance for each distribution? Calculate this by hand, not numerically.
c) The MATLAB function RAND(n,1) generates n samples from a uniformly distributed r.v. in the interval (0,1). Using this, generate n samples of the uniform p.d.f. in (1). Plot the sample p.d.f. for a number of different n's. What is a value of n that will produce a good discription of the true p.d.f.? Use hist(y,x) to produce a histogram of function y over the x-grid.
d) Repeat part (c) except for the normal distribution using the MATLAB function RANDN.

5. Consider the bivariate Gaussian p.d.f.:

$$p_{\rm xy}(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}\sigma_{\rm x}\sigma_{\rm y}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_x)^2}{\sigma_{\rm x}^2} - 2\rho\frac{(x-\mu_x)(y-\mu_y)}{\sigma_{\rm x}\sigma_{\rm y}} + \frac{(y-\mu_y)^2}{\sigma_{\rm y}^2}\right]\right\}.$$

a) Show that the marginal densities $p_{\rm x}(x)$ and $p_{\rm y}(y)$ are Gaussian. It is sufficient to show that $p_{\rm x}(x)$ is Gaussian and note the symmetric roles of x and y in the definition of $p_{\rm xy}(x,y)$.
b) Show that the conditional p.d.f. of $p_{\rm x|y}(x|y)$ is $\mathcal{N}(\mu_{x|y}, \sigma^2_{x|y})$ where $\mu_{x|y} = \mu_x + \rho\sigma_{\rm x}(y - \mu_y)/\sigma_{\rm y}$ and $\sigma^2_{x|y} = \sigma^2_x(1-\rho^2)$.