Problem Set 3

1. (Todling ch. 2, #5) Show that the following are admissible candidates for a covariance function:

   (a) $C_{\mathrm{x}}(\mathbf{r}) = a\delta(\mathbf{r})$, for $a > 0$ and $\mathbf{r} \in \mathcal{R}^n$,

   (b) $C_{\mathrm{x}}(x, y) = C_{\mathrm{x}}(r = |x - y|) = \Pi \exp(-r^2)$, for $x, y \in \mathcal{R}$. (Hint: In this case, the proof can be obtained by either showing that (4.23) is true or showing that (4.24) is satisfied. Use (4.23) and expand $\exp(2xy)$ in Taylor series.)

2. Let us return to the problem of ch. 3, #1, where we are lost at sea have no idea of our location. This time, we will solve this problem using a maximum likelihood method. We have two observations using two completely different instruments. For simplicity, we assume that the position is a 1D variable. The observation from the first instrument is z_1 with error variance σ_1^2. The observation from the second instrument is z_2 with error variance σ_2^2. The observation errors from the two instruments are uncorrelated and each is unbiased. There is no background information available. Additionally, assume z_1, z_2 and the state (position) x are all Gaussian. The measurement equation is: $\mathbf{z} = \mathbf{H}\mathrm{x} + \mathbf{v}$ where

$$\mathbf{z} = \begin{pmatrix} \mathrm{z}_1 \\ \mathrm{z}_2 \end{pmatrix}, \mathbf{v} = \begin{pmatrix} \mathrm{e}_1^r \\ \mathrm{e}_1^r \end{pmatrix}, \mathbf{H} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

   (a) What is the conditional probability of the observations, given the state, i.e. $p_{\mathbf{z}|\mathrm{x}}(\mathbf{z}|\mathrm{x})$? What is the solution that maximizes this conditional p.d.f.?

   (b) Now assume we have a background. With Gaussian error statistics, the MAP and MV estimators are the same and have analysis error covariance:

$$E[(\hat{\mathbf{x}}_{\mathrm{MAP}} - \mathbf{x})(\hat{\mathbf{x}}_{\mathrm{MAP}} - \mathbf{x})^{\mathrm{T}}] \quad = \quad \mathbf{P}_x = (\mathbf{P}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})^{-1} \tag{1}$$

   Note that the analysis error is independent of the observations, $\mathbf{z}$. Thus, an error analysis can be performed *before* the observations are taken, to decide how accurate the estimate will be when (or if) it is actually calculated. Assume that the prior is known to be $\mathcal{N}$(0,8), and that v_1 is $\mathcal{N}$(0,1) and v_2 is $\mathcal{N}$(0,s). s is a design parameter. System specifications require that the final estimate have variance less than or equal to 0.5. Since accurate meters cost money, it is reasonable to try to find the maximum value of s that is acceptable. Find this s.

3. Given the definition of the error estimate:

$$\mathbf{P}_x^{-1} = \mathbf{P}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}, \tag{2}$$

   show that the solution of J_{MAP} in

$$J_{\mathrm{MAP}}(\mathbf{x}) \quad = \quad (\mathbf{z} - \mathbf{H}\mathbf{x})^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{z} - \mathbf{H}\mathbf{x}) + (\mathbf{x} - \boldsymbol{\mu})^{\mathrm{T}}\mathbf{P}^{-1}(\mathbf{x} - \boldsymbol{\mu}) \tag{3}$$

is given by:

$$\hat{\mathbf{x}} = \mathbf{P}_x(\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{z} + \mathbf{P}^{-1}\boldsymbol{\mu}) \tag{4}$$

4. 1D Passive advection. Let us return to the problem of ch. 3, problem 6. We will repeat the analysis except with a 3DVAR instead of an OI scheme. Obtain var3d.m for this MATLAB exercise.

   (a) Compare var3d.m and oi.m. What are the relevant differences?

   (b) Under what conditions are OI and 3DVAR equivalent, in theory? Are these conditions satisfied in this problem?

   (c) Use the answer to problem 3 above to determine the analysis error covariance matrix for 3DVAR. Add this calculation to the code (1 statement). (Hint: How is $\mathbf{P}^a = \mathbf{P}_x$ related to J?)

   (d) Run the code. Type $\boxed{\text{var3d(0,1,0.95,0)}}$. The Courant number will be fixed at 0.95 and T_{final}=1. Three questions will be asked. Hitting "return" will give the default. First enter observation frequency of 5 (obs every 5 time steps), an observation sparsity of 1 (obs at every gridpoint), and "return" for the obs error standard deviation. This will give the default value of 0.02. Compare the speed of 3DVAR and OI. You can run OI by typing $\boxed{\text{oi(0,1,0.95,0)}}$. Note that a qualitative comparison is sufficient. An quantitative comparison is not possible because the 3dvar code includes the OI solution, for comparison purposes. Compare the solution to the OI case. Now uncomment the code that uses Newton's method and comment the call to *fminsearch*. Compare the speed and accuracy again to OI. Note that Newton's method converges in 1 iteration because our cost function is purely quadratic. Which is faster, Newton's method or *fminsearch*? Is Newton's method feasible for a large scale application?

   (e) Now let's see what happens when there are data gaps. Type $\boxed{\text{var3d(0,1,0.95,0)}}$, but answer "return" to all questions. This gives an obs every time step, over the left half of the domain with a std deviation of 0.02. How does the analysis fare? Now decrease the observation frequency by typing first 2 then 5 and keeping the obs pattern and error std dev the same as before. Now what happens to the solution? Why?

Newton method solution for a quadratic cost function

Consider a quadratic cost function:

$$J(\mathbf{x}) = \frac{1}{2}\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} + \mathbf{b}^{\mathrm{T}}\mathbf{x} + c \tag{5}$$

where $\mathbf{A}$ is symmetric, positive semi-definite. The gradient is

$$\nabla J(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}. \tag{6}$$

The Hessian or second derivative is

$$J''(\mathbf{x}) = \mathbf{A}. \tag{7}$$

Since we want to minimize (5), we want to solve for $\nabla J(\mathbf{x}) = \mathbf{0}$. Now for an initial guess $\mathbf{x}_g$, we have that

$$\nabla J(\mathbf{x}_g) = \mathbf{A}\mathbf{x}_g + \mathbf{b} \tag{8}$$

while at the solution, $\hat{\mathbf{x}}$, the gradient is zero, i.e.

$$\nabla J(\hat{\mathbf{x}}) = 0 = \mathbf{A}\hat{\mathbf{x}} + \mathbf{b}. \tag{9}$$

Subtracting (8) - (9) yields:

$$\nabla J(\mathbf{x}_g) = \mathbf{A}(\mathbf{x}_g - \hat{\mathbf{x}}). \tag{10}$$

Solving for $\hat{\mathbf{x}}$, we obtain

$$\hat{\mathbf{x}} = \mathbf{x}_g - \mathbf{A}^{-1}\nabla J(\mathbf{x}_g) \tag{11}$$

or on substituting the Hessian for $\mathbf{A}$:

$$\hat{\mathbf{x}} = \mathbf{x}_g - (J'')^{-1}\nabla J(\mathbf{x}_g). \tag{12}$$