ASE 381P Problem Set #4

Posting Date: October 10, 2019

You need not hand in anything. Instead, be prepared to answer any of these problems on an upcoming take-home exam. You may discuss your solutions with classmates up until the time that the exam becomes available, but do not swap work.

1. Recall that Newton's method for the general system of n nonlinear equations in n unknowns

$$0 = \boldsymbol{f}(\boldsymbol{x})$$

takes the iterative form

$$\boldsymbol{x}_g \leftarrow \boldsymbol{x}_g - \left[\frac{d\boldsymbol{f}}{d\boldsymbol{x}}|_{\boldsymbol{x}_g}\right]^{-1} \boldsymbol{f}(\boldsymbol{x}_g)$$

Use Newton's method to solve the following system of 2 nonlinear equations in 2 unknowns:

$$0 = x_1 + x_2 + x_1 x_2 + 5 \tag{1}$$
$$0 = (x_1)^2 + 2x_2 - (x_2)^2 - 2 \tag{2}$$

Start Newton's method from the following 3 first guesses: $\boldsymbol{x}_g = [4; -4]$, $\boldsymbol{x}_g = [6; 0]$, and $\boldsymbol{x}_g = [-5; 5]$. Record your first 6 Newton $\boldsymbol{x}_g$ iterates for each initial guess along with the value of the norm of $\boldsymbol{f}(\boldsymbol{x}_g)$ that goes with each iterate.

2. Prove that, under suitable assumptions, there exists a value of the Levenberg-Marquardt parameter λ that yields a cost decrease of the nonlinear least-squares cost function. In other words, prove that there is a $\lambda \geq 0$ such that $J(\boldsymbol{x}_g + \Delta\boldsymbol{x}) < J(\boldsymbol{x}_g)$ if

$$\Delta\boldsymbol{x} = (H^T H + \lambda I)^{-1} H^T [\boldsymbol{z} - \boldsymbol{h}(\boldsymbol{x}_g)]$$

Hints: In place of λ use $1/\epsilon$. Define a function $\tilde{J}(\epsilon) = J[\boldsymbol{x}_g + \Delta\boldsymbol{x}(\epsilon)]$. Compute the derivative of $\tilde{J}$ with respect to ϵ at $\epsilon = 0$, and make an argument about the sign of this derivative. Finish by using the sign of this derivative to prove that a cost decrease occurs for some positive value of ϵ.

3. Use the Gauss-Newton method to solve, in a weighted least-squares sense, the following over-determined system of nonlinear equations for $\boldsymbol{x} = [x_1; x_2; x_3]$:

$$z_j = x_1 \cos(x_2 t_j + x_3) + w_j \quad \text{for} \;\; j = 1, 2, ..., 11$$

where $E[w_j] = 0$, $E[w_j^2] = 1$, $E[w_i w_j] = 0.5$ if $|i - j| = 1$, $E[w_i w_j] = 0$ if $|i - j| \geq 2$. In other words, R is a matrix with 1s on its main diagonal, 0.5s on the diagonals just above and just below the main diagonal, and zeros everywhere else. The sample times t_j are contained in the vector

```
thist =[0; 0.1000; 0.2000; 0.3000; 0.4000; 0.5000;
        0.6000; 0.7000; 0.8000; 0.9000; 1.0000]
```

and the measurements z_j are contained in the vector

```
zhist = [7.7969; 1.4177; -3.0970; -7.6810; -9.8749; -6.1828;
         -0.8212; 4.5074; 8.2259; 9.5369; 6.2827]
```

Also, determine the estimation error covariance P_{xx} after you have minimized the appropriate weighted least-squares cost function (or the equivalent unweighted least-squares cost function). Make a record of your $\hat{\boldsymbol{x}}$ estimate and your P_{xx} estimation error covariance matrix.

Hints: You will note that the solution is sensitive to the initial guess $\boldsymbol{x}_g$. You'll want to try various initial guesses and compare their costs to eliminate local minima with high costs. To bound your search, note that the true solution has $0 < x_i < 20$ for all i.

4. Use the Gauss-Newton method to solve a modified missile tracking problem. The modified problem is like the problem presented in class, except that two radar stations are used instead of one, and each radar only measures range. Both radar stations are at zero altitude. Radar "a" is located at $y_1 = 4.1 \times 10^5$ m $= l_a$, and radar "b" is located at $y_1 = 4.4 \times 10^5$ m $= l_b$. Radar "a" has a range measurement error standard deviation of $\sigma_{\rho a} = 10$ m, but radar "b" has a measurement error standard deviation of $\sigma_{\rho b} = 30$ m because it is an older radar. Twelve samples of data are available from the two radar stations. The sample times are stored in the array `thist`, and the measurements are stored in the arrays `rhoahist` and `rhobhist`, all of which are contained in the attached MATLAB data file "radarmeasdata_missile.mat."

   Estimate the missile initial position and velocity using this data and determine the Gauss-Newton approximation of the estimation error covariance.

   Hints: Get your initial guess for the solution in the following manner: Assume perfect measurements and use two measurements that are far apart in time to derive positions for the missile at the corresponding times. (You may note that, because of noise, some measurements do not yield position solutions. Avoid these measurements.) Use those two positions to derive the corresponding initial conditions. You will have to solve two quadratic equations in two unknowns in order to go from two ranges to missile position coordinates. There is a clever trick that reduces these to one linear equation in one unknown. Assume that the missile's altitude is positive in order to resolve the ambiguity in the other unknown. Such ambiguity is inherent in any quadratic equation.

5. Consider the first-order scalar Markov process

$$y(k+1) = fy(k) + v(k), \quad k = 0, 1, ...$$

where $y(0) = 0$ and the elements of the process noise sequence $\{v(k)\}$ are independent and identically distributed as

$$v(k) \sim \mathcal{N}\left[0, (1-f^2)\sigma_{d_0}^2\right]$$

and where $f = e^{-T/\tau_{d_0}}$ for some constant time step T. One can think of τ_{d_0} as the process's decorrelation time, and of $\sigma_{d_0}^2$ as determining the process's intensity. For this process, do the following:

(a) Find the solution of $y(k+1)$ in terms of f and $\{v(j)\}_0^k$.

(b) Let $P_{yy}(k) \triangleq \mathbb{E}\left[\{y(k) - \bar{y}(k)\}^2\right]$, where $\bar{y}(k) \triangleq \mathbb{E}\left[y(k)\right]$. Find the steady state value

$$P_{yyss} \triangleq \lim_{k\to\infty} P_{yy}(k)$$

in terms of $\sigma_{d_0}^2$.

(c) Find a simple expression for $P_{yy}(k+1)$ in terms of $\sigma_{d_0}^2$ and f. Your expression should not be recursive nor should it include a summation. Plot the shape of of $P_{yy}(k+1)$ as a function of k.

(d) Let $R(k,j) \triangleq \mathbb{E}\left[y(k)y(j)\right]$. Find an expression for $R(k, k+n)$ in terms of f and $P_{yy}(k)$.

6. Consider the cost function

$$J(x) = \|z - h(x)\|^2$$

where $h(x) = \text{atan}(x)$ is the scalar measurement model. Let $x = 0$ be the true value of x and $z = 0$ be the corresponding noise-free measurement. Starting with an initial guess $\hat{x}^{(0)} = 1.5$, write a function that applies the Gauss-Newton procedure to solve for x^*, the value of x that minimizes $J(x)$. Write your code so that it implements step-size adjustment only when a configuration flag is asserted, otherwise setting the step size parameter $\alpha = 1$ for each iteration.

Make a plot of the cost $J(x)$ for $-6 \le x \le 6$ and overlay on this plot the trajectory of $J(\hat{x}^{(i)})$, for $i = 0, 1, 2, 3, 4$ and for both the step-size adjusted and non-step-size adjusted algorithms.

Qualitatively describe the cost trajectory with and without step-size adjustment. Is the crude step-size adjustment algorithm introduced in lecture adequate to avoid divergence of the Gauss-Newton algorithm in this case? Note that the crude adjustment algorithm makes no attempt to choose an optimal (efficient) α. Devise a technique for choosing the step size α more nearly *optimally* at each iteration.

Also do the following problems from Bar Shalom: 4-1.3 (prove by differentiating eq. 4.2.4-6), 4-3, 4-4 [define the effective window length as the length N_e of a sliding window average (which averages N_e previous values of $v(k)$) such that the sliding window average and $y(k)$ have the same variance as $k \to \infty$], 4-6, 4-7