Supplemental Exercises: Unit 3

Scientific Computing with Case Studies

Dianne P. O'Leary

SIAM Press, 2009

---

1. (a) If we use a quasi-Newton method to minimize a function, why is it important that the approximation to the Hessian be positive definite?

(b) Consider the DFP formula for approximating the Hessian inverse:

$$\boldsymbol{C}^{(k+1)} = \boldsymbol{C}^{(k)} - \frac{\boldsymbol{C}^{(k)}\boldsymbol{y}^{(k)}\boldsymbol{y}^{(k)T}\boldsymbol{C}^{(k)}}{\boldsymbol{y}^{(k)T}\boldsymbol{C}^{(k)}\boldsymbol{y}^{(k)}} + \frac{\boldsymbol{s}^{(k)}\boldsymbol{s}^{(k)T}}{\boldsymbol{y}^{(k)T}\boldsymbol{s}^{(k)}}$$

Compute (and simplify) $\boldsymbol{C}^{(k+1)}\boldsymbol{y}^{(k)}$.

---

2. Let

$$\hat{f}(\boldsymbol{x}) = e^{x_1+x_2} + x_1^2 + x_2^2 - x_1.$$

(a) (5) What are the necessary conditions that must be satisfied if a point $\hat{\boldsymbol{x}}$ is a local minimizer of this function $\hat{f}$?

(b) (5) Write a MATLAB program to apply 5 steps of Newton's method to approximately minimize $\hat{f}$ starting at $\boldsymbol{x}^{(0)} = [1, 0.3863]^T$. Use no linesearch; i.e., take a full step ($\alpha = 1$).

---

3. Consider the limited memory quasi-Newton method using the DFP update formula with $\boldsymbol{C}^{(0)} = \boldsymbol{I}$:

$$\boldsymbol{C}^{(k+1)} = \boldsymbol{C}^{(k)} - \frac{\boldsymbol{C}^{(k)}\boldsymbol{y}^{(k)}\boldsymbol{y}^{(k)T}\boldsymbol{C}^{(k)}}{\boldsymbol{y}^{(k)T}\boldsymbol{C}^{(k)}\boldsymbol{y}^{(k)}} + \frac{\boldsymbol{s}^{(k)}\boldsymbol{s}^{(k)T}}{\boldsymbol{y}^{(k)T}\boldsymbol{s}^{(k)}}$$

As an example, let $k = 2$.

(a) What vectors would you store in order to be able to form $\boldsymbol{C}^{(3)}\boldsymbol{v}$ for an arbitrary vector $\boldsymbol{v}$?

(b) How many floating-point multiplications would it take to form $\boldsymbol{C}^{(3)}\boldsymbol{v}$?

4. Suppose we measure $y(t_i)$, $i = 1, \ldots, 100$, and we model the relationship between t and y by

$$y_{pred}(t) = x_2 e^{x_1 t}$$

for some parameters x_1 and x_2. We want the "optimal" parameters, the values that minimize the least squares error:

$$\sum_{i=1}^{n} (y(t_i) - y_{pred}(t_i))^2.$$

Consider minimizing this function using `fmin`, a MATLAB-supplied function that minimizes a function of a **single** variable, or `fminunc`, a MATLAB-supplied function that minimizes a function of a vector of variables.

Write a MATLAB function `fcomp = f(x1)` that will evaluate the function to be minimized by `fmin`. (If you don't know how to do this, then for a maximum of 5 points, write a MATLAB function `fcomp = f(x)` to be used by `fminunc`.)