Math 171B: Numerical Optimization: Nonlinear Problems

Instructor: Michael Holst

Spring Quarter 2015

Solutions for Homework Assignment #3

Exercise 3.1. *Let H be a symmetric matrix with spectral decomposition $H = VDV^T$.*

(a) *Show that an eigenvector v associated with a positive eigenvalue λ satisfies $v^T Hv > 0$.*

By definition, $Hv = \lambda v$ so

$$v^T Hv = v^T \lambda v = \lambda v^T v = \lambda > 0$$

as $\lambda > 0$.

(b) *Write down the inverse of H in terms of V and D.*

We have $H^{-1} = (VDV^T)^{-1} = (V^T)^{-1}D^{-1}V^{-1}$. Since V is an orthogonal matrix, $VV^T = I$, so $V^T = V^{-1}$. So $H^{-1} = VD^{-1}V^T$.

(c) *If r is a positive integer, give an expression for H^r in terms of D and V. If H is positive definite, find a matrix B such that $H = B^2 = BB$ (B is the "square root" of H).*

To calculate H^r:

$$\begin{aligned}
H^r &= \overbrace{(VDV^T)(VDV^T)\cdots(VDV^T)}^{r \text{ times}} \\
&= VD(V^TV)DV^T\cdots(VDV^T) \\
&= VD(I)DV^T\cdots(VDV^T) \\
&= VD^2V^T\cdots(VDV^T) \\
&= \cdots = VD^rV^T.
\end{aligned}$$

Let $B = VD^{1/2}V^T$. Then $B^2 = VD^{1/2}V^TVD^{1/2}V^T = VD^{1/2}D^{1/2}V^T = VDV^T = H$.

(d) *Let α denote a scalar such that the matrix $H - \alpha I$ is nonsingular. If $\psi(\alpha)$ is the univariate function $\psi(\alpha) = u^T(H - \alpha I)^{-1}u$, where u is a nonzero vector, find $\psi'(\alpha)$.*

We have $\psi(\alpha) = u^T((VDV^T) - \alpha I)^{-1}u = u^T((VDV^T) - \alpha(VV^T))^{-1}u = u^T((V(D-\alpha)V^T)^{-1}u$. Using part (b):

$$\psi(\alpha) = u^T((V(D-\alpha)V^T)^{-1}u = u^T((V(D-\alpha)^{-1}V^T)u = W(D-\alpha)^{-1}W^T.$$

where $W = u^TV$. Then,

$$\psi(\alpha) = \sum_{i=1}^{n} \frac{w_i w_i}{\lambda_i - \alpha}$$

and

$$\begin{aligned}
\psi'(\alpha) &= \sum_{i=1}^{n} \frac{w_i w_i}{(\lambda_i - \alpha)^2} \\
&= W \begin{pmatrix} \frac{1}{(\lambda_1-\alpha)^2} & 0 & \cdots & 0 \\ 0 & \frac{1}{(\lambda_2-\alpha)^2} & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{(\lambda_n-\alpha)^2} \end{pmatrix} W^T
\end{aligned}$$

Exercise 3.2. *Given each of the following cases of a gradient $g(\bar{x})$ and Hessian $H(\bar{x})$ defined at a point $\bar{x}$, discuss the optimality of $\bar{x}$. (Do NOT use* MATLAB*. You may need to know how to compute eigenvalues by hand in your next exam).*

(i) $g(\bar{x}) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad H(\bar{x}) = \begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix}.$

Since $g(\bar{x}) \neq \vec{0}$, the first-order condition is not satisfied and $\bar{x}$ is not a local minimizer.

(ii) $g(\bar{x}) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad H(\bar{x}) = \begin{pmatrix} 3 & 2 \\ 2 & 0 \end{pmatrix}.$

The first-order condition is satisfied. However, since the eigenvalues of $H(\bar{x})$ are -1 and 4, $H(\bar{x})$ is indefinite, and the second-order condition fails. Thus $\bar{x}$ is not a local minimizer.

(iii) $g(\bar{x}) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad H(\bar{x}) = \begin{pmatrix} 4 & -1 \\ -1 & 4 \end{pmatrix}.$

The first-order condition is satisfied. The second-order sufficient condition is also satisfied as the eigenvalues of $H(\bar{x})$ are 5 and 3 (i.e. $H(\bar{x})$ is positive definite). Thus $\bar{x}$ is an isolated local minimizer.

(iv) $g(\bar{x}) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad H(\bar{x}) = \begin{pmatrix} -2 & 0 \\ 0 & -3 \end{pmatrix}.$

The first-order condition is satisfied. Since the eigenvalues of $H(\bar{x})$ are -2 and -3, $H(\bar{x})$ is negative definite, so $\bar{x}$ is an isolated local maximizer.

(v) $g(\bar{x}) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \qquad H(\bar{x}) = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$

The first-order condition is satisfied. Since the eigenvalues of $H(\bar{x})$ are 3, 2, and 0, $H(\bar{x})$ is positive semi-definite, and the second-order necessary condition is satisfied. However, $\bar{x}$ may not be a local minimizer.

Exercise 3.3.* *Write a* MATLAB *function with specification [f,g,H] = ex33(x) that computes $f(x), g(x), and H(x)$ for the function*

$$f(x) = e^{x_3}x_1^2 + 2x_2^2 + x_3^2 \cos x_1$$

at any point x. Use your function to compute $f(x), g(x)$, and $H(x)$ at $x = (0,0,0)^T$ and $x = (-1,2,-2)^T$. In each case, compute the spectral decomposition of the Hessian matrix and indicate if the necessary and sufficient conditions for unconstrained local minimization are satisfied.

See the TA for the solution. (Note: The first part is similar to Exercise 1.6, and the second is similar to the exercise above.)

Exercise 3.4. *Let $q(x) \in \mathbb{R}^n$ be the quadratic function $q(x) = c^T x + \frac{1}{2}x^T Hx$, where H is symmetric.*

(a) *Write down an expression for $\nabla q(x)$ in terms of c, H, and x.*

$\nabla q(x) = c + Hx$. (Note: This is similar to Exercise 1.5 (b).)

(b) *Given an arbitrary point x_0 and a direction p, write down the Taylor-series expansion of $q(x_0 + p)$.*

Using part (a), and noting that $q''(x_0) = H$, the Taylor-series expansion of $q(x_0 + p)$ is

$$\begin{aligned} q(x_0 + p) &= q(x_0) + \nabla q(x_0)^T p + \frac{1}{2}p^T q''(x_0)p \\ &= q(x_0) + (c + Hx_0)^T p + \frac{1}{2}p^T Hp. \end{aligned}$$

(c) *For this part, consider $q(x)$ such that H is positive definite. If p is a direction such that $\nabla q(x_0)^T p < 0$, show that there exists a positive minimizer α^* of $q(x_0 + \alpha p)$. Derive a closed-form expression for α^*.*

If α^* is to be a minimizer, then $\nabla q(x_0 + \alpha^* p) = 0$ (note that since H is positive definite, second-order conditions are already satisfied). Given H positive definite and p such that $\nabla q(x_0)^T p < 0$, we want to show α^* is positive. Consider $\nabla q(x_0 + \alpha^* p)^T p$:

$$
\begin{aligned}
0 = \nabla q(x_0 + \alpha^* p)^T p &= (c + H(x_0 + \alpha^* p))^T p \\
&= c^T p + (x_0 + \alpha^* p)^T H p
\end{aligned}
$$

since $H = H^T$. So

$$
\begin{aligned}
\alpha^* &= \frac{-(c^T + x_0^T H)p}{p^T H p} \\
&= \frac{-\nabla q(x_0)^T p}{p^T H p}.
\end{aligned}
$$

Given that H is positive definite, the denominator of the above expression is positive. Also, since $\nabla q(x_0)^T p < 0$, the numerator is also positive, so α^* is positive.

Exercise 3.5.* *Write a* MATLAB *m-file steepest.m that implements the method of steepest descent with a backtracking line search. Your function must include the following features.*

- *Use $\mu = \frac{1}{4}$ to define the sufficient-decrease criterion in the backtracking algorithm.*
- *The minimization must be terminated when either $||g(x_k)|| \le 10^{-5}$ or 75 iterations are performed. Any* MATLAB *"while" loop must include a test that will terminate the loop if it is executed more than 20 times.*

Use steepest.m to find a minimizer of the function

$$f(x) = e^{x_3} x_1^2 + 2x_2^2 + x_3^2 \cos x_1,$$

starting at the point $(-1, 1, 1)^T$. Next, minimize the function (first write a MATLAB *function as in Exercise 3.3)*

$$f(x) = x_1 + x_2 + x_3 + x_4 + x_1^2 + x_2^2 + 10^{-1} x_3^2 + 10^{-3} x_4^2,$$

starting at the point $(-1, 0, 1, 1)^T$. Compare the two runs. Can you explain why steepest descent behaves like this?

See TA for questions about the code. The error constant for this method depends on the ratio of the largest and smallest eigenvalues of the Hessian at x^*. Since the second function has $\text{cond}(H) = \lambda_{\max}/\lambda_{\min} = 2/.002 = 1000$, the second function has an ill-conditioned Hessian. See pages 88-89 for more info.