TMA4180 Optimization Theory
Spring 2013

Norwegian University of Science and Technology
Department of Mathematical Sciences

Exercise set 2

Tutorial: Thursday 31.01 16:15-17:00 in Kjl 4.

1 Let M be a real $n \times n$ matrix, i.e. $M = (m_{i,j})_{i,j=0}^n, m_{i,j} \in \mathbb{R}$.

- M is "symmetric" iff $M = M^{\mathrm{T}}$.
- M is "positive definite" iff $\forall x \in \mathbb{R}^n \setminus \{0\} : x^{\mathrm{T}} M x > 0$
- M is "positive semi-definite" iff $\forall x \in \mathbb{R}^n : x^{\mathrm{T}} M x \geq 0$
- M is called "diagonally dominant" iff all diagonal entries $m_{i,i}$ are greater than 0 and

$$\forall i = 1, \dots, n : \qquad |m_{i,i}| \geq \sum_{j \neq i} |m_{i,j}|.$$

It is known that all eigenvalues of a symmetric matrix are real. Furthermore, every real symmetric matrix M can be diagonalized as

$$M = Q^{\mathrm{T}} D Q,$$

where Q is an orthogonal matrix (i.e., $Q^{\mathrm{T}} = Q^{-1}$) and $D = diag(\lambda_1, \lambda_2, \dots, \lambda_n)$ is a diagonal matrix whose diagonal entries correspond to the eigenvalues of M.

a) Let $u \in \mathbb{R}^n$. Prove that the $n \times n$ matrix $A := uu^{\mathrm{T}}$ is symmetric and positive semi-definite.
Solution: Symmetric:

$$(uu^{\mathrm{T}})^{\mathrm{T}} = (u^{\mathrm{T}})^{\mathrm{T}} u^{\mathrm{T}} = uu^{\mathrm{T}}.$$

Positive-semidefinite:

$$\forall x \in \mathbb{R}^n : \qquad x^{\mathrm{T}}(uu^{\mathrm{T}})x = (x^{\mathrm{T}}u)(u^{\mathrm{T}}x) = (x^{\mathrm{T}}u)^2 \geq 0$$

b) Prove: The real symmetric matrix $M \in \mathbb{R}^{n \times n}$ is positive semi-definite if and only if all its eigenvalues λ_i are greater than or equal to 0.
Solution: Let $x \in \mathbb{R}^n$. We can diagonalize M as $Q^{\mathrm{T}} D Q$ (see above):

$$x^{\mathrm{T}} M x \geq 0 \Leftrightarrow x^{\mathrm{T}} Q^{\mathrm{T}} D Q x \geq 0 \Leftrightarrow (Qx)^{\mathrm{T}} D (Qx) \geq 0.$$

The matrix Q is orthogonal, so this is equivalent to:

$$x^{\mathrm{T}} D x \geq 0.$$

The diagonal matrix D consists of all of M's eigenvalues $\{\lambda_i\}_{i=1}^n$ in the diagonal; picking for x the canonical basis-vectors $e_1, e_2, \dots e_n$ of $\mathbb{R}^n$ shows that all λ_i must be greater than or equal to zero.

c) Prove: A real symmetric, diagonally dominant matrix M is positive semi-definite. *(Hint: Use Gershgorin's circle theorem: Let $M \in \mathbb{R}^{n \times n}$. For every row i, let $R_i := \sum_{j \neq i} |m_{i,j}|$. Denote by $D_c(r)$ the closed disc of radius r centered at c. Then: Every eigenvalue of M lies in at least one of the discs $D_{m_{i,i}}(R_i)$).*

Solution: Denote the eigenvalues of M by $\lambda_1, \dots, \lambda_n$. We know from Gershgorin's theorem that:

$$\forall i \in \{1, \dots, n\} \; \exists j \in \{1, \dots, n\} : \lambda_i \in D_{m_{j,j}}(R_j).$$

This means that:

$$m_{j,j} + R_j \geq \lambda_i \geq m_{j,j} - R_j \geq 0,$$

where the last inequality follows from M's diagonal dominance. The conclusion then follows from b).

2

a) Find the gradient $\nabla f(x)$ and the Hessian $\nabla^2 f(x)$ of the following functions:

$$f(x) = x_1^2 - 5x_1x_2 + x_2^4 - 25x_1 - 8x_2 \tag{1a}$$

$$f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2 \tag{1b}$$

$$f(x) = e^{x_1}(4x_1^2 + 2x_2^2 + 4x_1x_2 + 2x_2 + 1) \tag{1c}$$

Solution: Elementary calculations.

b) Do one iteration with the steepest descent algorithm on (1.a), starting from $(0, 0)$. Repeat the exercise using Newton's method.
Solution: Elementary calculations. Compare with the numerical results in c) and d).

c) The enclosed MATLAB file `sd.m` is a rather primitive implementation of the steepest descent (SD) algorithm for solving the minimization problem

$$\min_{x \in \mathbb{R}^2} f(x)$$

numerically, with (1.a) as the test problem. In `sd_wp.m` the same algorithm is implemented, for a 2-dimensional problem it gives a graphical presentation of the iterations. As stopping criteria for the iterations, we have used $\|\nabla f_k\| \leq tol$, with $tol = 1.e - 4$.

Use one of the codes to solve the problem (1.a). Try with different initial values, and comment on what you observe. Solve the two other problems as well.

d) Implement and test Newton's method ((2.15) in N&W) on the problems above. How many iterations are needed now? Compare with the steepest descent method.
Solution: See `newton.m`.

e) Solve the problem(s) by the use of MATLAB's `fminsearch`. Set the option `Display` to `on`, so you can see the output of each iterations.

It may be useful to read description of the algorithm in the documentation.

3 Consider the problem

$$\min_{x \in \mathbb{R}^n} f(x), \qquad \text{where } f(x) = \frac{1}{2}x^{\mathrm{T}}Ax - b^{\mathrm{T}}x,$$

and A is a positive definite matrix. Show that if we start in a point x_0 where gradient vector g_0 $(= \nabla f(x_0))$ is an eigenvector for A, then the steepest descent method converges in only one step.

Solution: Assume $Ag_0 = \lambda g_0$. We already know that

$$\alpha^* = \frac{\|g_0\|^2}{\|g_0\|_A^2} = \frac{g_0^{\mathrm{T}} g_0}{g_0^{\mathrm{T}} A g_0} = \frac{1}{\lambda}.$$

and

$$x_1 = x_0 - \frac{1}{\lambda} g_0.$$

Now,

$$Ax_1 = Ax_0 - \frac{1}{\lambda}(\lambda g_0) = Ax_0 - (Ax_0 - b) = b,$$

which shows that $x_1 = A^{-1}b$ is the minimum.