Instructor: Emre Mengi

Math 409/509 (Spring 2011)

Study Guide for Homework 1
This homework concerns the topics listed below. Please don't hesitate to ask for help if any of these topics is unclear. Section numbers refer to the sections of the class-text by Philip Gill and Margaret Wright.

- Problem formulation (Section 1.1)
- The gradient vector and Hessian matrix (Section 1.7.1 - see Definition 1.7.2 and 1.7.3)
- Vector norms (Section 1.3.1)
- The derivatives of vector-valued functions, in particular the Jacobian matrix (Section 1.7)
- The chain rule for vector-valued functions (Section 1.7.2)
- Taylor's theorem (Section 1.6 and 1.7.3)
- Local and global minimizers (Section 3.1)
- Optimality conditions for unconstrained optimization (with derivations)
  - First order necessary conditions
  - Second order necessary conditions
  - Second order sufficient conditions

  (Section 3.2; see also Theorem 2.2, 2.3 and 2.4 on pages 14-16 in Nocedal and Wright)
- Eigenvalues, eigenvectors and singular values
  - Basic Properties
  - Calculation by hand

  (Section 1.2)
- Positive definite, positive semidefinite matrices and their eigenvalue characterizations (Section 1.2)

Homework 1 (due on March 7th, Monday by 2pm)
Most of the questions in this homework do not require computations in Matlab. (In the other homeworks you will have plenty of Matlab questions.) Only in questions 2.c attach the Matlab output to your answer. Please return the solutions to only seven out of the nine questions below.

1. Compute the gradient and Hessian for each of the following functions.

(i) $f(x_1, x_2) = e^{x_2} x_1^2 + \sin(x_1) x_2^2$ where $x_1, x_2 \in \mathbb{R}$.

(ii) $f : \mathbb{R}^n \to \mathbb{R}$, $f(x) = x^T x$.

(iii) $f : \mathbb{R}^n \to \mathbb{R}$, $f(x) = (x^T x) e^{x^T x}$.

(iii) $f : \mathbb{R}^n \to \mathbb{R}$, $f(x) = b^T x$ where $b \in \mathbb{R}^n$.

(v) $f : \mathbb{R}^n \to \mathbb{R}$ defined as

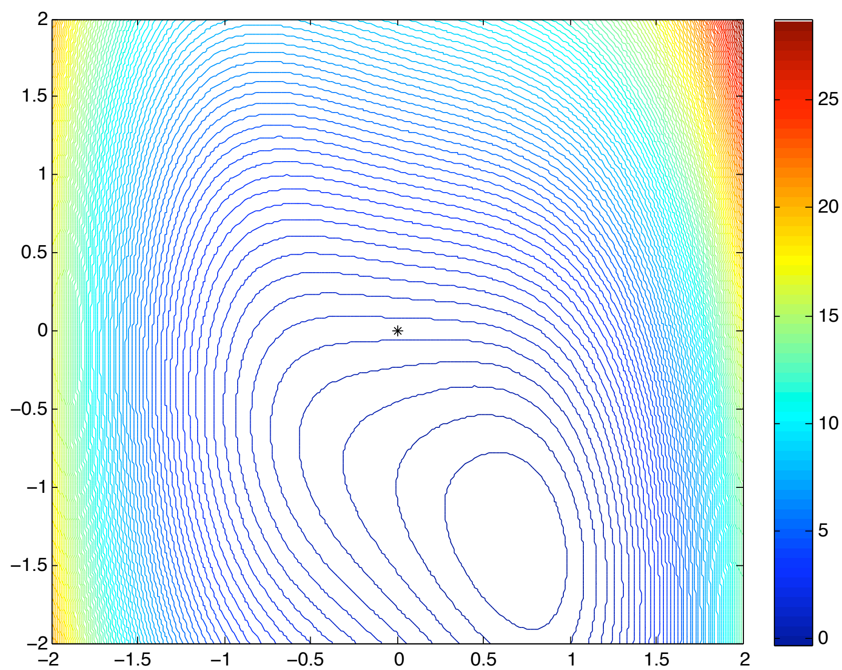
$$f(x) = \frac{1}{2} x^T A x + b^T x + c$$

where $b \in \mathbb{R}^n$, A is an $n \times n$ matrix and c is a real scalar.

2. (Modified from Practical Methods of Optimization, Fletcher)
Obtain expressions for all first and second derivatives of the function of two variables

$$f(x) = x_1^4 + x_1 x_2 + (1 + x_2)^2$$

for which a contour diagram is provided below.



(a) Argue why the point $(0, 0)$ marked with an asterisk on the contour diagram cannot be a local minimizer.

(b) Show that the Hessian $\nabla^2 f(0)$ does not satisfy the property $\underline{p^T \nabla^2 f(0) p > 0 \text{ for all } p \neq 0}$.

(c) A local minimizer of f is $x_* = (0.6959, -1.3479)$. Verify that the first order necessary conditions for optimality are satisfied at x_*. (Note: Use Matlab for this verification.)

3. Suppose the displacement of an object is measured one and two seconds after it starts moving as 10 and 50 meters, respectively. You aim to fit a function of the form

$$\phi_{(c_1, c_2)}(t) = c_1 e^{c_2 t}$$

that best captures the displacement of the object based on the measurements at $t = 1$ and $t = 2$.

(a) The objective function that you are trying to minimize is the sum of the squares of the modeling errors at $t = 1$ and $t = 2$. In general we define the modeling error at time t as

$$r_{(c_1,c_2)}(t) = \left|y_t - \phi_{(c_1,c_2)}(t)\right|$$

where y_t is the measured displacement at time t. Pose this problem as an unconstrained optimization problem over the variables c_1 and c_2.

(b) Write down the first order necessary conditions (that a local minimizer must satisfy) for the unconstrained optimization problem in **(a)** in terms of the variables c_1 and c_2.

4. The ∞-norm $\|\cdot\|_\infty$ and 1-norm $\|\cdot\|_1$ for vectors are defined as

$$\|v\|_\infty = \max_{1\leq i\leq n} |v_i|$$

and

$$\|v\|_1 = \sum_{i=1}^{n} |v_i|$$

(a) Prove that $\|v\|_\infty$ is a vector norm.

(b) Prove that $\|v\|_1$ is a vector norm.

(c) Describe in words the sets given below.

$$\begin{aligned} \mathcal{S}_2^\infty &= \{v \in \mathbb{R}^2 : \|v\|_\infty = 1\}, & \mathcal{S}_3^\infty &= \{v \in \mathbb{R}^3 : \|v\|_\infty = 1\} \\ \mathcal{S}_2^1 &= \{v \in \mathbb{R}^2 : \|v\|_1 = 1\}, & \mathcal{S}_3^1 &= \{v \in \mathbb{R}^3 : \|v\|_1 = 1\} \end{aligned}$$

5. Given the function $f : \mathbb{R}^2 \to \mathbb{R}$,

$$f(x) = [x_1 \;\; x_2] \begin{bmatrix} 2 & -1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + [4 \;\; -2] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

and the vector $p = [1 \;\; -1]^T$. Let $\ell : \mathbb{R} \to \mathbb{R}$, $\ell(\alpha) = f(x+\alpha p)$ and $L : \mathbb{R} \to \mathbb{R}^2$, $L(\alpha) = \nabla f(x+\alpha p)$. Compute the derivatives, $\ell'(\alpha)$, $\ell''(\alpha)$ and $L'(\alpha)$.

6. By Taylor's theorem a univariate function $f : \mathbb{R} \to \mathbb{R}$ that is twice continuously differentiable has the quadratic expansion

$$f(x+h) = f(x) + f'(x)h + \frac{1}{2}f''(x+th)h^2$$

for some $t \in (0, 1)$. Prove the generalization of Taylor's theorem for a multivariate function, that is given a function $f : \mathbb{R}^n \to \mathbb{R}$ that is twice continuously differentiable and a vector $p \in \mathbb{R}^n$ show that there exists a $t \in (0, 1)$ such that

$$f(x+p) = f(x) + \nabla f(x)^T p + \frac{1}{2}p^T \nabla^2 f(x+tp)p.$$

(Hint : Apply Taylor's theorem for a univariate function to $\ell(\alpha) = f(x + \alpha p)$.)

7. Answer all parts for each of the quadratic polynomials

$$\begin{aligned} f(x_1, x_2) = &\quad x_1^2 + 4x_2^2 - 4x_1x_2 + x_1 + x_2 + 4 \\ g(x_1, x_2) = &\quad 2x_1^2 + 3x_2^2 + 2x_1x_2 + 4x_1 - x_2 + 3 \\ h(x_1, x_2, x_3) = &\quad x_1^2 - 2x_2^2 + x_3^2 + 3x_1 + 4x_2 - x_3 + 8. \end{aligned}$$

(a) Write the polynomial in the form

$$\frac{1}{2}x^T Ax + b^T x + c$$

where A is a symmetric matrix, x is the vector of variables, b is a vector and c is a scalar.

(b) Does the polynomial have a *stationary point* where the gradient is zero? If it has, find the stationary point.

(c) If a stationary point of the polynomial exists, is the stationary point a local minimizer? Justify your answer.

8. (Practical Methods Of Optimization, Flecther; Numerical Optimization, Nocedal and Wright) Show that each of the functions

$$\begin{aligned} f(x) &= (x_2 - x_1^2)^2 + x_1^5 \\ g(x) &= 100(x_2 - x_1^2)^2 + (1 - x_1)^2 \end{aligned}$$

has a unique stationary point and determine in each case whether the unique stationary point is a local minimizer or not.

9. Let $f : \mathbb{R}^n \to \mathbb{R}$ be twice continuously differentiable. Suppose that there exist $x, p \in \mathbb{R}^n$ such that $f(x) = f(x+p)$ and $\nabla^2 f(x+\alpha p)$ is *positive definite* for all $\alpha \in (0, 1)$. Show that some $t \in (0, 1)$ is a local minimizer of $\ell(\alpha) = f(x + \alpha p)$. Does this imply that the point $x + tp$ is a local minimizer of f?