Optimisation methods – Homework 1

Claudio Maggioni

March 21, 2021

1 Exercise 1

1.1 Gradient and Hessian

The gradient and the Hessian for f are the following:

$$\nabla f = \begin{bmatrix} 2x_1 + x_2 \cdot \cos(x_1) \\ 9x_2^2 + \sin(x_1) \end{bmatrix}$$

$$H_f = \begin{bmatrix} 2 - x_2 \cdot \sin(x_1) & \cos(x_1) \\ \cos(x_1) & 18x_2 \end{bmatrix}$$

1.2 Taylor expansion

$$f(h) = 0 + \langle \begin{bmatrix} 0+0 \\ 0+0 \end{bmatrix}, \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} \rangle + \frac{1}{2} \langle \begin{bmatrix} 2-0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix}, \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} \rangle + O(\|h\|^3)$$

$$f(h) = \frac{1}{2} \langle \begin{bmatrix} 2h_1 + h_2 \\ h_1 \end{bmatrix}, \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} \rangle + O(\|h\|^3)$$

$$f(h) = \frac{1}{2} \left(2h_1^2 + 2h_1 h_2 \right) + O(\|h\|^3)$$

$$f(h) = h_1^2 + h_1 h_2 + O(\|h\|^3)$$

2 Exercise 2

2.1 Gradient and Hessian

For A symmetric, we have:

$$\frac{d}{dx}\langle b, x\rangle = \langle b, \cdot\rangle = b$$
$$\frac{d}{dx}\langle Ax, x\rangle = 2\langle Ax, \cdot\rangle = 2Ax$$

Then:

$$\nabla J = Ax - b$$
$$H_J = \frac{d}{dx}\nabla J = A$$

2.2 First order necessary condition

It is a necessary condition for a minimizer x^* of J that:

$$\nabla J(x^*) = 0 \Leftrightarrow Ax^* = b$$

2.3 Second order necessary condition

It is a necessary condition for a minimizer x^* of J that:

$$\nabla^2 J(x^*) \geq 0 \Leftrightarrow A \text{ is positive semi-definite}$$

2.4 Sufficient conditions

It is a sufficient condition for x^* to be a minimizer of J that the first necessary condition is true and that:

$$\nabla^2 J(x^*) > 0 \Leftrightarrow A \text{ is positive definite}$$

2.5 Does $\min_{x\in R^n} J(x)$ have a unique solution?

Not in general. If for example we consider A and b to be only zeros, then $J(x) = 0$ for all $x \in R^n$ and thus J would have an infinite number of minimizers.

However, for if A would be guaranteed to have full rank, the minimizer would be unique because the first order necessary condition would hold only for one value x^*. This is because the linear system $Ax^* = b$ would have one and only one solution (due to A being full rank).

3 Exercise 3

3.1 Quadratic form

$f(x, y)$ can be written in quadratic form in the following way:

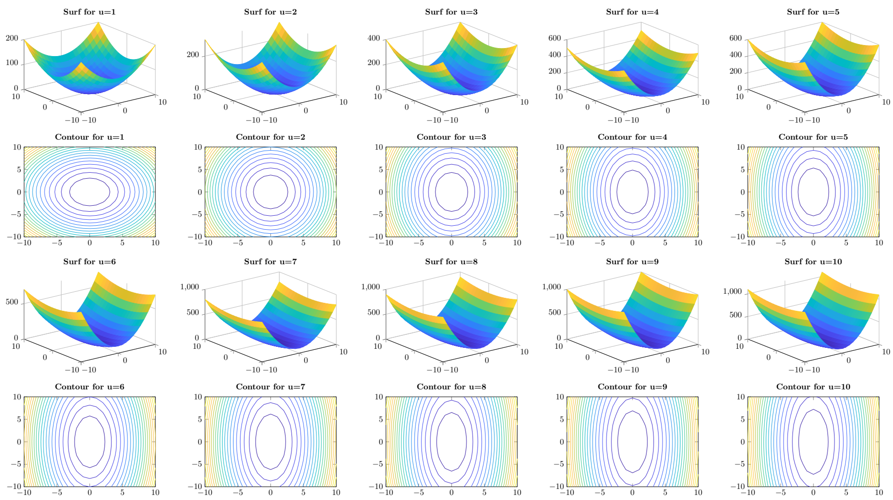
$$f(v) = \frac{1}{2}\langle \begin{bmatrix} 2 & 0 \\ 0 & 2\mu \end{bmatrix} v, v\rangle + \langle \begin{bmatrix} 0 \\ 0 \end{bmatrix}, x\rangle$$

where:

$$v = \begin{bmatrix} x \\ y \end{bmatrix}$$

3.2 Matlab implementation with `surf` and `contour`

The graphs generated by MATLAB are shown below:



Isolines get stretched along the y axis as μ increases. For a large μ, points well far away from the axes could be a problem since picking search directions and steps using a naive gradient based method iterations will zig-zag to the minimizer reaching it slowly.