Numerical Optimization

Shuai Lu 170742

Homework 1

Exercise 1.1
(a) [2 point] Let A $\in R^{m\times n}$. Prove that

$$kerA := \{v \in R^n : Av = 0\}$$

and

$$RangeA := \{w \in R^m : \exists v \in R^n : Av = w\}$$

are subspaces of R^n and R^m, respectively.

Solution:
$\exists v = 0 \in R^n, st. Av = 0$. So $v \in kerA, kerA \neq \phi$.
$\forall \alpha, \beta \in R$, if $x, y \in kerA, A(\alpha x + \beta y) = \alpha Ax + \beta Ay = 0$. So $\alpha x + \beta y \in kerA$.
So $kerA$ is a linear subspace of R^n.

$v = 0 \in R^n, st. w = Av = 0 \in R^m$. So $w \in RangeA, RangeA \neq \phi$.
$\forall \alpha, \beta \in R$, if $x, y \in RangeA$ and $x = Av_1, y = Av_2$, $\alpha x + \beta y = \alpha Av_1 + \beta Av_2 = A(\alpha v_1 + \beta v_2)$. So $\alpha x + \beta y \in RangeA$.
So $RangeA$ is a linear subspace of R^m.
(b) [1 point] Compute the rank of

$$A = \begin{bmatrix} 1 & 2 & 3 \\ -3 & 4 & 1 \\ 1 & -1 & 0 \end{bmatrix}$$

.

Solution:

Using Gaussian Elimination

$$A = \begin{bmatrix} 1 & 2 & 3 \\ -3 & 4 & 1 \\ 1 & -1 & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 2 & 3 \\ 0 & 10 & 10 \\ 0 & -3 & -3 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 2 & 3 \\ 0 & 10 & 10 \\ 0 & 0 & 0 \end{bmatrix}$$

It is obvious that $RankA = 2$.

Exercise 1.2
Consider $f : R^2 \to R$,

$$f(x, y) = x^2 \cos y$$

(a) [1 point] Write down its gradient $\nabla f(x, y)$. Evaluate this gradient at $(x, y) = \left(3, \frac{\pi}{4}\right)$.

Solution:

$$\nabla f = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{pmatrix} = \begin{pmatrix} 2x \cos y \\ -x^2 \sin y \end{pmatrix}. \text{ So } \nabla f|_{(3, \frac{\pi}{4})} = \begin{pmatrix} 3\sqrt{2} \\ -\frac{9\sqrt{2}}{2} \end{pmatrix}.$$

(b) [1 point] Write down its Hessian $\nabla^2 f(x, y)$. Due to which properties of f, the Hessian is symmetric? Evaluate this Hessian at $(x, y) = \left(3, \frac{\pi}{4}\right)$. Is the Hessian at this point positive or negative (semi-) definite, or indefinite?

Solution:
$$\nabla^2 f = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix} = \begin{pmatrix} 2\cos y & -2x \sin y \\ -2x \sin y & -x^2 \cos y \end{pmatrix}.$$

The Hessian is symmetric because the second partial derivatives of f are continuous.
$\nabla^2 f\big|_{(3,\frac{\pi}{4})} = \begin{pmatrix} \sqrt{2} & -3\sqrt{2} \\ -3\sqrt{2} & \frac{-9\sqrt{2}}{2} \end{pmatrix} = A.$
Solving $det(\lambda I - A) = 0$, we get the $\lambda_1 \lambda_2 = \frac{-27}{2} < 0$. So the Hessian at $(x, y) = \left(3, \frac{\pi}{4}\right)$ is indefinite.

Exercise 1.3
(a) [4 points] Implement a program (in C/C++/Java) that solves the linear system

$$Ax = b$$

for a given matrix $A \in R^{n\times n}$ and a given vector $b \in R^n (n < 100)$ by the Gaussian elimination. For the associated LU-decomposition, print the upper triangular matrix U and the product $L^{-1}b$. Test your code for

$$A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 3 & -2 \\ 1 & 1 & 2 \end{bmatrix}, \quad b = \begin{bmatrix} 5 \\ 4 \\ 6 \end{bmatrix}$$

(print the solution x).

Solution:
As shown in Fig1.

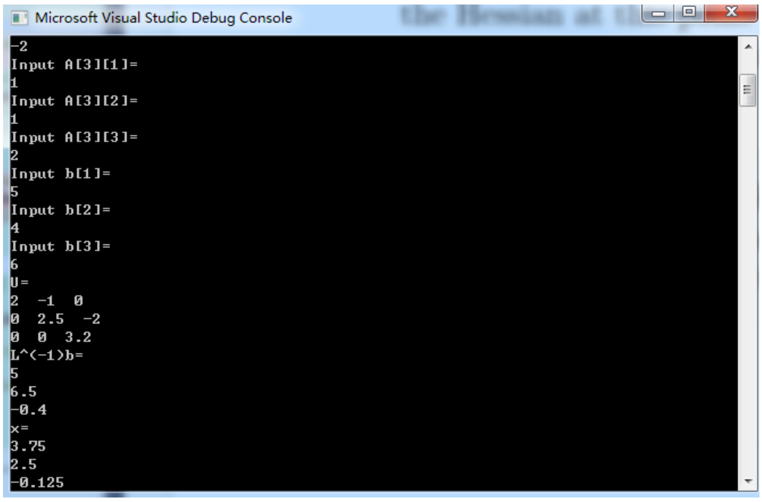

Figure 1: Gaussian Elimination

(b) [3 points] Extend your implementation with the pivoting strategy that uses the entry with the greatest absolute value in the column as the pivoting element. Test your code for

$$A = \begin{bmatrix} 2 & -2 & 0 \\ -2 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}, \quad b = \begin{bmatrix} 5 \\ 4 \\ 6 \end{bmatrix}$$

(print the upper triangular matrix U and the solution x).

Solution:
As shown in Fig2.

Exercise 1.4
Consider a two times continuously differentiable function $f : R^n \to R$.
(a) [2 points] Let $x^* \in R^n$ be a local minimizer of f. Prove that $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*) \geq 0$.

Solution:
$x^* \in R^n$ is a local minimizer of f, so $\exists \delta > 0$, if $\|x - x^*\| < \delta$, st. $f(x) - f(x^*) \geq 0$
Using Taylor expansion: $f(x) - f(x^*) = \nabla f(x^*)^{\mathrm{T}} (x - x^*) + o(x - x^*)^2 \geq 0$.
Let $x = \alpha \nabla f(x^*) + x^*$, $o(x - x^*)^2$ is neglectable when α is small enough. So $\nabla f(x^*)^{\mathrm{T}} (x - x^*) = \alpha \nabla f(x^*)^{\mathrm{T}} \nabla f(x^*) \geq$

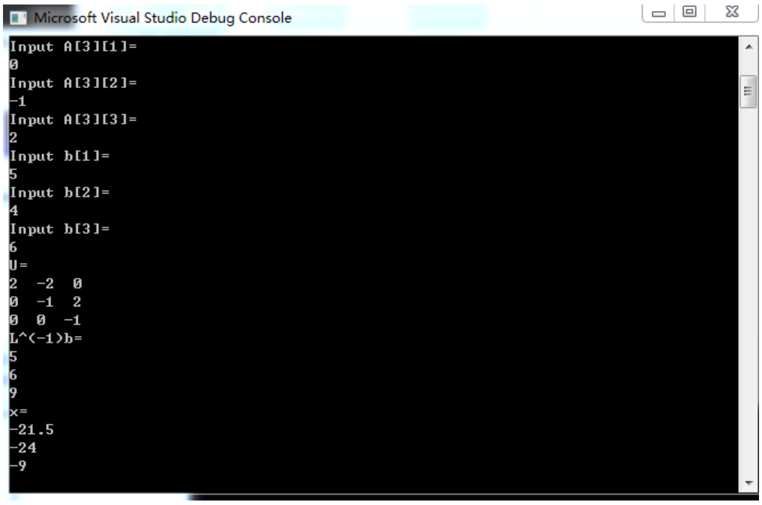

Figure 2: Gaussian Elimination with Pivoting

0.
Because α can be negative or positive, so $\nabla f\left(x^*\right)^{\mathrm{T}} \nabla f\left(x^*\right) = 0$, which means $\nabla f\left(x^*\right) = 0$.

(b) [2 points] Let $x^* \in R^n$ be a point with $\nabla f\left(x^*\right) = 0$ and $\nabla^2 f\left(x^*\right) > 0$. Prove that x^* is a local minimizer of f. Demonstrate by an example for $n = 1$ that in this statement, the condition for $\nabla^2 f$ connot be weakened to $\nabla^2 f \geq 0$.

Solution:
Using Taylor expansion: $f(x) - f(x^*) = \nabla f\left(x^*\right)^{\mathrm{T}}\left(x - x^*\right) + \frac{1}{2}(x - x^*)^{\mathrm{T}} \nabla^2 f\left(x^*\right)\left(x - x^*\right) + o(x - x^*)^3$.
$\exists \delta > 0$ small enough, when $\|x - x^*\| < \delta$, we can neglect $o(x - x^*)^3$. Due to $\nabla f\left(x^*\right) = 0$ and $\nabla^2 f\left(x^*\right) > 0$.
So $f(x) - f(x^*) = \frac{1}{2}(x - x^*)^{\mathrm{T}} \nabla^2 f\left(x^*\right)\left(x - x^*\right) \geq 0$, when $\|x - x^*\| < \delta$. So x^* is a local minimizer of f.
When $n = 1$, $f(x) = x^3$ st. $\nabla f(x)|_{x=0} = 0$, $\nabla^2 f\big|_{x=0} = 0 \geq 0$, but $x = 0$ is not a local minimizer of $f(x) = x^3$.