

Numerical Optimization Assignment 2

103062512 徐丞裕

1.

(a) Decompose A using LDL^T decomposition:

$$\begin{aligned}
 A &= \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix} \begin{bmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{bmatrix} \begin{bmatrix} 1 & l_{21} & l_{31} \\ 0 & 1 & l_{32} \\ 0 & 0 & 1 \end{bmatrix} \\
 &= \begin{bmatrix} d_1 & d_1 l_{21} & d_1 l_{31} \\ d_1 l_{21} & d_2 + d_1 l_{21}^2 & d_2 l_{32} + d_1 l_{21} l_{31} \\ d_1 l_{31} & d_2 l_{32} + d_1 l_{21} l_{31} & d_3 + d_1 l_{31}^2 + d_2 l_{32}^2 \end{bmatrix} \\
 &= \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}
 \end{aligned}$$

(b) Compute the eigenvalues of A :

$$\begin{aligned}
 Av &= \lambda v \\
 (A - \lambda I)v &= 0
 \end{aligned}$$

Since the eigenvector v is not a zero vector, the system above has non-trivial solutions, which implies $\det(A - \lambda I) = 0$.

$$\det(A - \lambda I) = 0$$

$$\det \left(\begin{bmatrix} 1 - \lambda & 2 & 1 \\ 2 & 3 - \lambda & 0 \\ 1 & 0 & -2 - \lambda \end{bmatrix} \right) = 0$$

$$(1 - \lambda)(3 - \lambda)(-2 - \lambda) - 4(-2 - \lambda) - 1(3 - \lambda) = 0$$

$$\lambda^3 - 2\lambda^2 - 10\lambda + 1 = 0$$

By solving the equation above, we then have:

$$\begin{aligned}
 \lambda &\approx 4.2812 \\
 \lambda &\approx 0.0982 \\
 \lambda &\approx -2.3794
 \end{aligned}$$

2.

(a)

$$\begin{aligned}
 \|Qx\| &= \sqrt{(Qx)^T(Qx)} \\
 &= \sqrt{x^T Q^T Q x} \\
 &= \sqrt{x^T x} \\
 &= \|x\|
 \end{aligned}$$

- (b) Since A is symmetric, we have $A = Q\Lambda Q^\top$ be the eigen decomposition of A where Q is an orthogonal matrix and $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_i)$.

$$\begin{aligned}
\|A\| &= \max_{\|x\|=1} \|Ax\| \\
&= \max_{\|x\|=1} \sqrt{(Ax)^\top (Ax)} \\
&= \max_{\|x\|=1} \sqrt{x^\top A^\top A x} \\
&= \max_{\|x\|=1} \sqrt{x^\top (Q\Lambda Q^\top)^\top (Q\Lambda Q^\top) x} \\
&= \max_{\|x\|=1} \sqrt{x^\top Q\Lambda^\top Q^\top Q\Lambda Q^\top x} \\
&= \max_{\|x\|=1} \sqrt{x^\top Q\Lambda^2 Q^\top x}
\end{aligned}$$

Furthermore, as Q is an orthogonal matrix (rows and columns are orthogonal unit vectors), we can rewrite the equation above as:

$$\begin{aligned}
\|A\| &= \max_{\|x\|=1} \sqrt{x^\top Q\Lambda^2 Q^\top x} \\
&= \max_{\|w\|=1} \sqrt{w^\top \Lambda^2 w}
\end{aligned}$$

where $w = Q^\top x$ is also a unit vector.

In order to maximize the function above, one should choose w with 1 at k^{th} element and 0 otherwise, where $\Lambda_{ii}^2 = \lambda_{max}^2$.

Thus we have $\|A\| = \sqrt{\lambda_{max}^2} = |\lambda_{max}|$.

3.

- (a) Gradient of $f(x, y)$:

$$\nabla f(x, y) = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} = \begin{bmatrix} 400x^3 - 400xy + 2x - 2 \\ 200(y - x^2) \end{bmatrix}$$

Hessian of $f(x, y)$:

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 1200x^2 - 400y + 2 & -400x \\ -400x & 200 \end{bmatrix}$$

- (b) Instead of implementing a function with other functions as its parameter, I used OOP paradigm and

- i. make the Rosenbrock's function a class with the following class methods:
 - A. `valueAt`: evaluates and returns the function value given a point.
 - B. `gradientVectorAt`: evaluates and returns the gradient vector given a point.
 - C. `hessianMatrixAt`: evaluates and returns the Hessian matrix given a point.
- ii. make the backtracking line search method a class with:
 - A. `BacktrackingLineSearcher`: a constructor, creates and returns an instance of the class with configurable parameters.
 - B. `fitStepLength`: finds and returns a suitable step length using the parameters configured at creation.

- (c) The following figure shows the changes of the norms gradient vectors with different algorithms (in 300 iterations even if the algorithm has reached the convergence criteria).

