Due date of the homework: February 3rd, 2023, midnight.

1. Compute the gradient and the Hessian of this function

$$f(x) = f(x_1, x_2, x_3) = 3x_1^2x_3 + 2x_3 - 4x_3x_2^2$$

at this operating point $x^{(0)} = [3 \ -1 \ 2]^\top$.

Gradient:

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} \\ \frac{\partial f(x)}{\partial x_2} \\ \frac{\partial f(x)}{\partial x_3}) \end{bmatrix} = \begin{bmatrix} 6x_1x_3 \\ -8x_3x_2 \\ 3x_1^2 - 4x_2^2 + 2 \end{bmatrix} \ \nabla f(x_0) = \begin{bmatrix} 36 \\ 16 \\ 25 \end{bmatrix}$$

Hessian:

$$\nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 x_2} & \frac{\partial^2 f(x)}{\partial x_1 x_3} \\ \frac{\partial^2 f(x)}{\partial x_2 x_1} & \frac{\partial^2 f(x)}{\partial x_2^2} & \frac{\partial^2 f(x)}{\partial x_2 x_3} \\ \frac{\partial^2 f(x)}{\partial x_3 x_1} & \frac{\partial^2 f(x)}{\partial x_3 x_2} & \frac{\partial^2 f(x)}{\partial x_3^2} \end{bmatrix} = \begin{bmatrix} 6x_3 & 0 & 6x_1 \\ 0 & -8x_3 & -8x_2 \\ 6x_1 & -8x_2 & 0 \end{bmatrix} \ \nabla^2 f(x_0) = \begin{bmatrix} 12 & 0 & 18 \\ 0 & -16 & 8 \\ 18 & 8 & 0 \end{bmatrix}$$

2. Compute the gradient and the Hessian of this function

$$f(x) = f(x_1, x_2) = 3x_1^2 cos(e^{-x_2}) + 2x_2^2 + 7x_1 - 8x_2 + x_1^2 + 4$$

at this operating point $x^{(0)} = [0 \ -1]^\top$.

Gradient:

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} \\ \frac{\partial f(x)}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 6x_1 cos(e^{-x_2}) + 7 + 2x_1 \\ 3x_1^2 sin(e^{-x_2})e^{-x_2} + 4x_2 - 8 \end{bmatrix} \ \nabla f(x_0) = \begin{bmatrix} 7 \\ -12 \end{bmatrix}$$

Hessian:

$$\nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 x_2} \\ \frac{\partial^2 f(x)}{\partial x_2 x_1} & \frac{\partial^2 f(x)}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} 6cos(e^{-x_2}) + 2 & 6x_1 sin(e^{-x_2})e^{-x_2} \\ 6x_1 sin(e^{-x_2})e^{-x_2} & -3x_1^2 e^{-x_2}(cos(e^{-x_2})e^{-x_2} + sin(e^{-x_2})) + 4 \end{bmatrix}$$

$$\nabla^2 f(x_0) = \begin{bmatrix} 6cos(e) + 2 & 0 \\ 0 & 4 \end{bmatrix}$$

3. For Problems 1 and 2, do the following:
   - Compute the first order and second (quadratic) order Taylor series approximation around the given operating points.
   - Using Matlab, plot the functions and their corresponding approximations.
   - Evaluate the definiteness of the quadratic approximations (i.e., are the resulting quadratic approximation positive definite, positive semidefinite, negative definite, etc...).

Problem 1:

$$\begin{aligned} f_{lin}(x) &= f(x_0) + \nabla f(x_0)^T (x - x_0) \\ &= 50 + \begin{bmatrix} 36 & 16 & 25 \end{bmatrix} \begin{bmatrix} x_1 - 3 \\ x_2 + 1 \\ x_3 - 2 \end{bmatrix} \\ &= 36x_1 + 16x_2 + 25x_3 - 92 \end{aligned}$$

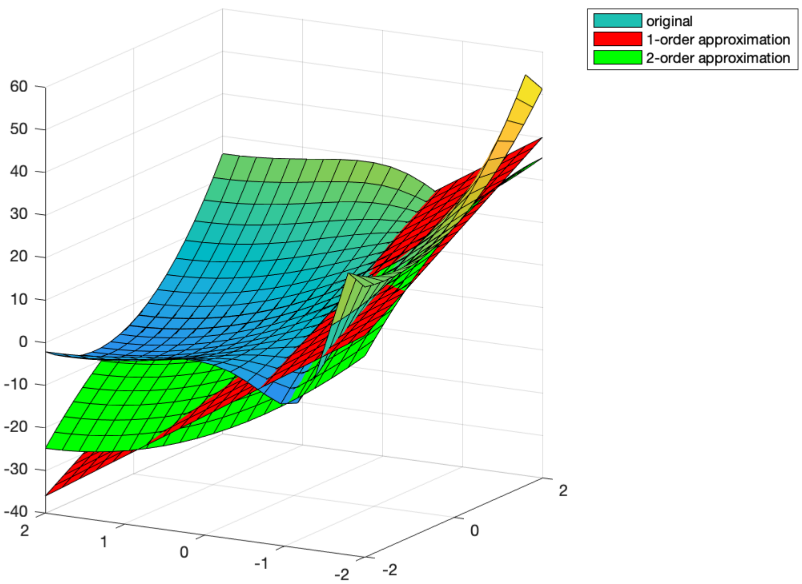

Figure 1: Function visualizations for problem 2

$$\begin{aligned}
f_{quad}(x) &= f(x_0) + \nabla f(x_0)^T(x-x_0) + \frac{1}{2}(x-x_0)^T\nabla^2 f(x_0)(x-x_0) \\
&= 50 + \begin{bmatrix}36 & 16 & 25\end{bmatrix}\begin{bmatrix}x_1-3\\x_2+1\\x_3-2\end{bmatrix} + \frac{1}{2}\begin{bmatrix}x_1-3 & x_2+1 & x_3-2\end{bmatrix}\begin{bmatrix}12 & 0 & 18\\0 & -16 & 8\\18 & 8 & 0\end{bmatrix}\begin{bmatrix}x_1-3\\x_2+1\\x_3-2\end{bmatrix} \\
&= -36x_1 - 16x_2 - 21x_3 + 6x_1^2 - 8x_2^2 + 18x_1x_3 + 8x_2x_3 + 46
\end{aligned}$$

The Hessian matrix of the quadratic approximation is $\begin{bmatrix}12 & 0 & 18\\0 & -16 & 8\\18 & 8 & 0\end{bmatrix}$, the eigenvalues are $-21.44, -8.07, 25.52$ and thus it is indefinite.

Problem 2:

$$\begin{aligned}
f_{lin}(x) &= f(x_0) + \nabla f(x_0)^T(x-x_0) \\
&= 14 + \begin{bmatrix}7 & -12\end{bmatrix}\begin{bmatrix}x_1\\x_2+1\end{bmatrix} \\
&= 7x_1 - 12x_2 + 2
\end{aligned}$$

$$\begin{aligned}
f_{quad}(x) &= f(x_0) + \nabla f(x_0)^T(x-x_0) + \frac{1}{2}(x-x_0)^T\nabla^2 f(x_0)(x-x_0) \\
&= 14 + \begin{bmatrix}7 & -12\end{bmatrix}\begin{bmatrix}x_1\\x_2+1\end{bmatrix} + \frac{1}{2}\begin{bmatrix}x_1 & x_2+1\end{bmatrix}\begin{bmatrix}6cos(e)+2 & 0\\0 & 4\end{bmatrix}\begin{bmatrix}x_1\\x_2+1\end{bmatrix} \\
&= -1.7x_1^2 + 2x_2^2 + 7x_1 - 8x_2 + 4
\end{aligned}$$

The Hessian matrix of the quadratic approximation is $\begin{bmatrix}6cos(e)+2 & 0\\0 & 4\end{bmatrix}$, the eigenvalues are $-3.47, 4$ and thus it is indefinite.

4. Compute the eigenvalues, eigenvectors, determinant, and rank of these two matrices

$$A = \begin{bmatrix} 1 & \pi \\ -1 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}.$$

Eigenvalues of matrix A: $\lambda_1 = \frac{5}{2} + \frac{i}{2}\sqrt{4\pi - 9}, \lambda_2 = \frac{5}{2} - \frac{i}{2}\sqrt{4\pi - 9}$

Eigenvectors of matrix A: $x_1 = \begin{bmatrix} \frac{3}{2} - \frac{i}{2}\sqrt{4\pi - 9} \\ 1 \end{bmatrix}, x_2 = \begin{bmatrix} \frac{3}{2} + \frac{i}{2}\sqrt{4\pi - 9} \\ 1 \end{bmatrix}$

Determinant of A: $4 + \pi$

Rank of A: 2 (full)

Eigenvalues of matrix B: $\lambda_1 = \frac{15}{2} + \frac{3\sqrt{33}}{2}, \lambda_2 = \frac{15}{2} - \frac{3\sqrt{33}}{2}, \lambda_3 = 0$

Eigenvectors of matrix B: $x_1 = \begin{bmatrix} -\frac{1}{2} + \frac{3}{22}\sqrt{33} \\ \frac{1}{4} + \frac{3}{44}\sqrt{33} \\ 1 \end{bmatrix}, x_2 = \begin{bmatrix} -\frac{1}{2} - \frac{3}{22}\sqrt{33} \\ \frac{1}{4} - \frac{3}{44}\sqrt{33} \\ 1 \end{bmatrix}, x_3 = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$

Determinant of B: 0

Rank of B: 2

The computations should be by hand but please verify your answers via Matlab. We did not cover that in Module 0, so feel free to use online content to guide you through how matrix properties are manually computed.

5. Verify that the one-norm given by

$$||x||_1 = \sum_{i=1}^{n} |x_i|$$

is indeed a norm on $\mathbb{R}^n$ via proving that it satisfies the three vector norm properties.

Positive definiteness: it is obvious that $||x||_1 = \sum_{i=1}^n |x_i| >= 0$, and $\sum_{i=1}^n |x_i| = 0$ if and only if $x_i = 0 \forall i$, i.e., $x = 0$

Scaling: suppose $a \in \mathbb{R}^n$, $||ax||_1 = \sum_{i=1}^n |ax_i| = \sum_{i=1}^n |a||x_i| = |a| \sum_{i=1}^n |x_i| = |a|||x||_1$

Triangle inequality: suppose $y \in \mathbb{R}^n$, $||x+y||_1 = \sum_{i=1}^n |x_i + y_i|$, it is obvious that $|x_i + y_i| <= |x_i| + |y_i|$ for $\forall i$ and the equality holds if and only if both x_i and y_i is nonnegative. Therefore, we have $||x+y||_1 = \sum_{i=1}^n |x_i + y_i| <= \sum_{i=1}^n (|x_i| + |y_i|) = ||x||_1 + ||y||_1$

6. Prove that for all vectors $x \in \mathbb{R}^n$, we have

$$||x||_1 \geq ||x||_2.$$

$||x||_1^2 = (\sum_{i=1}^n |x_i|)^2 = \sum_{i=1}^n |x_i|^2 + 2\sum_{i \neq j} |x_i x_j| \geq \sum_{i=1}^n |x_i|^2 = (\sqrt{\sum_{i=1}^n x_i^2})^2 = ||x||_2^2$, then we have $||x||_1 \geq ||x||_2$. since both $||x||_1$ and $||x||_2$ are nonnegative.

7. Compute $||A||_F$, $||A||_2$, $||A||_1$ and $||A||_\infty$ of this matrix:

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

$||A||_F = \sqrt{trace(A*A)} = \sqrt{trace(\begin{bmatrix} 5 & -4 & 1 \\ -4 & 6 & 0 \\ 1 & 0 & 5 \end{bmatrix})} = 4$

$||A||_2 = \sqrt{\lambda_{max}(A*A)} = \sqrt{\lambda_{max}(\begin{bmatrix} 5 & -4 & 1 \\ -4 & 6 & 0 \\ 1 & 0 & 5 \end{bmatrix})} = \sqrt{9.6272} = 3.1028$

$||A||_1 = max_{1\leq j\leq n} \sum_{i=1}^{m} |a_{ij}| = 4$

$||A||_\infty = max_{1\leq i\leq m} \sum_{j=1}^{n} |a_{ij}| = 4$

8. Prove that this new norm for any square matrix A with dimension n defined as

$$||A||_{**} = \max|a_{ij}|, \quad \text{for all } 1 \leq i, j \leq n$$

is not a legit norm. You can find a counter example that contradicts one of the basic matrix norm properties.

It does not satisfy sub-multiplicative property: suppose we have two matrix $A = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$, then $||AB||_{**} = ||\begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}||_{**} = 2 >= ||A||_{**}||B||_{**} = 1$

9. Prove that the Frobenius norm of any matrix A is indeed a legitimate matrix norm by showing that it satisfies the basic matrix norm properties.

Positive definiteness: $||A||_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2} >= 0$, the equality holds if and only if $a_{ij} = 0$ for $\forall i, j$

Homogeneity: for any scalar α, we have $||\alpha A||_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|\alpha a_{ij}|^2} = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|\alpha|^2|a_{ij}|^2} = |\alpha|\sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2} = |\alpha|||A||_F$

Sub-multiplicative property: suppose B is another m by n matrix. $||AB||_F^2 = \sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}b_{ij}|^2 <= \sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}b_{ij}|^2 + \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k,l\neq i,j}|a_{ij}b_{kl}|^2 = (\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2) * (\sum_{i=1}^{m}\sum_{j=1}^{n}|b_{ij}|^2) = ||A||_F^2||B||_F^2$, therefore $||AB||_F <= ||A||_F||B||_F$

Sub-additive property: suppose B is another m by n matrix. $||A+B||_F^2 = \sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij} + b_{ij}|^2 <= \sum_{i=1}^{m}\sum_{j=1}^{n}(|a_{ij}| + |b_{ij}|)^2 = \sum_{i=1}^{m}\sum_{j=1}^{n}(|a_{ij}|^2 + |b_{ij}|^2 + 2|a_{ij}||b_{ij}|) = ||A||_F^2 + ||B||_F^2 + 2||AB||_F^2 <= ||A||_F^2 + ||B||_F^2 + 2||A||_F^2||B||_F^2 = (||A||_F + ||B||_F)^2$, therefore $||A+B||_F <= ||A||_F + ||B||_F$

10. Find the *full* and *thin* singular value decompositions for these two matrices by hand and verify them via Matlab

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

Slide 44/50 in Module 0 explains how to compute these decompositions.

$AA^T = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$, we have eigenvalues of $\lambda_1 = 3, \lambda_2 = 1$ and eigenvectors of $x_1 = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} \end{bmatrix}, x_2 = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}$, thus $U = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}, \Sigma = \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$

$A^T A = \begin{bmatrix} 2 & -1 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$, we have eigenvalues of $\lambda_1 = 3, \lambda_2 = 1, \lambda_3 = 0$ and eigenvectors of $x_1 = \begin{bmatrix} \frac{\sqrt{6}}{3} \\ -\frac{\sqrt{6}}{6} \\ \frac{\sqrt{6}}{6} \end{bmatrix}, x_2 = \begin{bmatrix} 0 \\ \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}, x_3 = \begin{bmatrix} \frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \\ -\frac{\sqrt{3}}{3} \end{bmatrix}$, thus $V = \begin{bmatrix} \frac{\sqrt{6}}{3} & 0 & \frac{\sqrt{3}}{3} \\ -\frac{\sqrt{6}}{6} & \frac{\sqrt{2}}{2} & \frac{\sqrt{3}}{3} \\ \frac{\sqrt{6}}{6} & \frac{\sqrt{2}}{2} & -\frac{\sqrt{3}}{3} \end{bmatrix}$

Thus the full SVD is $A = U\Sigma V^T = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{6}}{3} & 0 & \frac{\sqrt{3}}{3} \\ -\frac{\sqrt{6}}{6} & \frac{\sqrt{2}}{2} & \frac{\sqrt{3}}{3} \\ \frac{\sqrt{6}}{6} & \frac{\sqrt{2}}{2} & -\frac{\sqrt{3}}{3} \end{bmatrix}^T$,

the thin SVD is $A = U_1 \Sigma_1 V_1^T = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{6}}{3} & 0 \\ -\frac{\sqrt{6}}{6} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{6}}{6} & \frac{\sqrt{2}}{2} \end{bmatrix}^T$

Similarly, for matrix B, we have $B = U\Sigma V^T$, where $U = \begin{bmatrix} \frac{1}{\sqrt{14}} & -\frac{2}{\sqrt{5}} & -\frac{3}{\sqrt{70}} \\ \frac{2}{\sqrt{14}} & \frac{1}{\sqrt{5}} & -\frac{6}{\sqrt{70}} \\ \frac{3}{\sqrt{14}} & 0 & \frac{5}{\sqrt{70}} \end{bmatrix}, \Sigma = \begin{bmatrix} \sqrt{70} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, V = \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix}$, and $B = U_1 \Sigma_1 V_1^T$, where $U_1 = \begin{bmatrix} \frac{1}{\sqrt{14}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{14}} & \frac{1}{\sqrt{5}} \\ \frac{3}{\sqrt{14}} & 0 \end{bmatrix}$, $\Sigma_1 = \begin{bmatrix} \sqrt{70} & 0 \\ 0 & 0 \end{bmatrix}, V_1 = \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix}$

11. Compute the 2- and nuclear-norms of matrices A and B given in the previous problems (by utilizing the results of your SVD).

2-norm of A: $||A||_2 = \sigma_{max} A = \sqrt{3}$
Nuclear-norm of A: $||A||_* = \sum_{i=1}^{min\{m,n\}} \sigma_i(A) = 1 + \sqrt{3}$
2-norm of B: $||B||_2 = \sigma_{max} B = \sqrt{70}$
Nuclear-norm of B: $||B||_* = \sum_{i=1}^{min\{m,n\}} \sigma_i(B) = \sqrt{70}$