In-class Exercise 2.3
Gradient Descent for $Ax = b$
AMATH 301
University of Washington
Jakob Kotas

1. The example from the 2.3 video was:

$$\underbrace{\begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix}}_{A}\underbrace{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}}_{x} = \underbrace{\begin{bmatrix} 2 \\ -8 \end{bmatrix}}_{b}$$

   (a) Write out the quadratic form $f(x) = \frac{1}{2}x^T Ax - b^T x$ and assume $c = 0$.

$$f(x) = \frac{1}{2}\begin{bmatrix} x_1 & x_2 \end{bmatrix}\begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} 2 & -8 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= \frac{3}{2}x_1^2 + 2x_1x_2 + 3x_2^2 - 2x_1 + 8x_2$$

   (b) Find the gradient $\nabla f(x)$.

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 3x_1 + 2x_2 - 2 \\ 2x_1 + 6x_2 + 8 \end{bmatrix}$$

   Visualization of #1

2. The code from the 2.3 video is hard-coded for the problem indicated in #1. Rewrite it so that $f(x)$, $\partial f/\partial x_1$, and $\partial f/\partial x_2$ are separate Python functions. (You can still assume that x is 2×1.)

   Then use this new code to approximate the solution to

$$\underbrace{\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}}_{A}\underbrace{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}}_{x} = \underbrace{\begin{bmatrix} 7 \\ 2 \end{bmatrix}}_{b}$$

   using Gradient Descent.

3. Use the code from #2 to approximate the solution to the following systems using Gradient Descent. What goes wrong in each case?

   (a)

$$\underbrace{\begin{bmatrix} 3 & -2 \\ -2 & 1 \end{bmatrix}}_{A}\underbrace{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}}_{x} = \underbrace{\begin{bmatrix} 11 \\ -6 \end{bmatrix}}_{b}$$

   Visualization of #3a

   (b)

$$\underbrace{\begin{bmatrix} 3 & -2 \\ 2 & 1 \end{bmatrix}}_{A}\underbrace{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}}_{x} = \underbrace{\begin{bmatrix} -12 \\ -1 \end{bmatrix}}_{b}$$