ECE 805 – Machine Learning

Tutorial 1, Friday, March 18, 2022

Norms:

- **Definition:** Let (X, F) be a linear vector space. A real-valued function is called a norm (and is denoted by || . ||) if the following properties hold:
  (i) $||x|| \geq 0$ and $||x||=0 \Rightarrow x=0_x$ for each $x \in X$
  (ii) $||\alpha \cdot x|| = |a| \cdot ||x||$ for each $x \in X$ and $\alpha \in F$
  (iii) $||x + y|| \leq ||x|| + ||y||$ for each $x, y \in X$ (triangle inequality)

- **Examples of norms:**
  1. $X = \Re^n$ (X is a vector with n elements)
     1-norm: $\|x\|_1 = \sum_{i=1}^{n} |x_i|$
     2-norm or Euclidean norm: $\|x\|_2 = \sqrt[2]{\sum_{i=1}^{n} |x_i|^2}$
     ∞-norm: $\|x\|_\infty = max_{1 \leq i \leq n} |x_i|$
     p-norm: $\|x\|_p = \sqrt[p]{\sum_{i=1}^{n} |x_i|^p}$

  2. X = C [0, T] (X is a continuous function in the given interval [0, T])
     1-norm: $\|x\|_1 = \int_0^T |x(t)|\, dt$
     2-norm or Euclidean norm: $\|x\|_2 = \sqrt[2]{\int_0^T |x(t)|^2\, dt}$
     ∞-norm: $\|x\|_\infty = max_{t \in [0,T]} |x(t)|$
     p-norm: $\|x\|_p = \sqrt[p]{\int_0^T |x(t)|^p\, dt}$

  3. $X \in \Re^{n \times n}$ (Matrix norms)
     1-norm: $\|X\|_1 = max_{1 \leq j \leq n} \sum_{i=1}^{n} |x_{ij}|$
     2-norm or Euclidean norm: $\|X\|_2 = \sqrt[2]{\lambda_{max}\, (A^T A)}$
     ∞-norm: $\|X\|_\infty = max_{1 \leq i \leq n} \sum_{j=1}^{n} |x_{ij}|$

Gradient Method – Steepest Descent Method: $\begin{bmatrix} x_{k+1} \end{bmatrix} = \begin{bmatrix} x_k \end{bmatrix} - \lambda_k^* \begin{bmatrix} \nabla F\, (x_k) \end{bmatrix}$

Newton's Method: $\begin{bmatrix} x_{k+1} \end{bmatrix} = \begin{bmatrix} x_k \end{bmatrix} - H(x_k)^{-1} \begin{bmatrix} \nabla F\, (x_k) \end{bmatrix}$

Exercises:

1. Compute the 1-norm, ∞-norm and Euclidean norm (2-norm) of the following matrices:

   (a) $A_1 = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ (b) $A_2 = \begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix}$

2. Using the first order necessary condition find the minimum points of the following functions:

   (a) $f(x) = x_1^2 + x_2^2 - x_1$

   (b) $f(x) = x_1^4 + x_1 x_2 + \frac{1}{2} x_2^2$

   (c) $f(x) = x_1^2 + 2x_2^2 + 4x_1 + 4x_2$

   (d) $f(x) = 4x_1^2 - 2x_1^4 + \frac{1}{3} x_1^6 + x_1 x_2 + \frac{1}{4} x_2^2$

   Verify that the points you find are minimums or not by checking the second order sufficient conditions.

3. Find the point on the line 3x +2y = 5 in two-dimensional space closest to the origin when distance is measured by each of the following three norms:
   a. The 1-norm
   b. The 2-norm
   c. The ∞-norm

4. (a) Let $x = [x_1 \;\; x_2 \;\; x_3]^T$ and $F(x) = x_1 x_2^2 + x_2 x_3^2 + x_3^3$. Compute the Gradient and Hessian of F and find all values x^* for which $\nabla F(x^*) = 0$.
   Is the Hessian singular or nonsingular at these values?

   (b) Using the first order necessary conditions ($\nabla F(x^*) = 0$) find a minimum point of the function:
   $F(x_1, x_2, x_3) = 2x_1^2 + x_1 x_2 + x_2^2 + x_2 x_3 + x_3^2 - 6x_1 - 7x_2 - 8x_3 + 9.$
   Verify that this point is a minimum by checking the second order sufficiency conditions.

5. Write a script to implement the Gradient method – Steepest Descent Method for minimizing the function:
   $F(x_1, x_2) = e^{x_1}(4x_1^2 + 2x_2^2 + 4x_1 x_2 + 2x_2 + 1).$
   Let your initial estimate be something close to the origin. Choose the step-size λ to be a constant. Run a few simulations with different values of λ to see what happens as you vary the step-size from a small to a large value.