The Exercise of Chap 1

Yanwei Fu

September 16, 2018

In this course we will heavily use vectors and matrices, as well as linear and quadratic functions. To refresh the memories of people who haven't used these in a long time, the rst part of this question makes you review basic operations on vectors and matrices. If you are rusty on basic vector and matrix operations, see the notes on linear algebra.

1 Vectors, Matrices, and Quadratic Functions

1.1 Basic Operations

Using the definitions below,

$\alpha = 10$, $x = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$,$y = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$,$A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$,

evaluate the following expressions (show your work):

1. $\alpha (x + y)$ (vector addition and scalar multiplication)
2. $x^T y + \|x\|^2$ (inner product and norm)
3. Ax (matrix-vector multiplication)
4. $A^T A$ (matrix-matrix multiplication)
5. $Tr\left(A^T A\right)$ (matrix trace)

If $\{x, y, z\}$ are real-valued column-vectors of length d and $\{A, B, C\}$ are real-valued d by d matrices, state whether each of the below statements is true or false in general; if the statements is false, please give the correct form.

1. $x^T (y + z) = z^T x + y^T x$
2. $x^T x = x x^T$
3. $x^T A y = y^T A^T x$
4. $x^T \left(y^T z\right) = \left(x^T y\right)^T z$
5. $AB = BA$

6. $A\left(B^TC\right) = \left(AB^T\right)C$
7. $A^T\left(B+C\right) = A^TB + A^TC$
8. $\left(A+B\right)^T = A^T + B^T$
9. $\left(AB\right)^T = A^TB^T$

1.2 Gradients and Hessians of Linear and Quadratic Functions

For this question we'll use the convention that all values are real and:

1. α is a scalar.
2. a and b are length-d column-vectors.
3. Element i of b is denoted by b_i.
4. A and B are d by d matrices.
5. Row i of A is denoted by a_i^T .
6. W is a *symmetric* d by d matrix.

Express the gradient $\nabla f\left(x\right)$ and Hessian $\nabla^2 f\left(x\right)$ of the following linear/quadratic functions in matrix notation, simplifying as much as possible.

1. $f\left(x\right) = a^Tx + \alpha$ (linear)
2. $f\left(x\right) = x^Ta + a^TAx + x^TA^Tb$ (more linear forms)
3. $f\left(x\right) = x^Tx + x^TWx + x^TABx$ (quadratic forms)
4. $f\left(x\right) = \frac{1}{2}\left(Ax-b\right)^TW\left(Ax-b\right)$ (weighted least squares)
5. $f\left(x\right) = \frac{\lambda}{2}\left\|x\right\|^2 + \frac{1}{2}\sum_{i=1}^{n}\left(a_i^Tx - b_i\right)^2$ (L2-regularized least squares)