The Exercise of Week 1

Xuening Zhu

September 14, 2019

We will heavily use vectors, matrices and quadratic functions in our course. The rest part of this question enables you to review the related knowledge.

1 Basic Operation

using the definitions,$\alpha = 8$, $x = \begin{bmatrix}1\\3\end{bmatrix}$ $y = \begin{bmatrix}2\\4\end{bmatrix}$ $A = \begin{bmatrix}1 & 3\\3 & 2\end{bmatrix}$ to evaluate the following expressions(show your work):

1. $\alpha(x + y)$
2. $x^T y + \|x\|^2$
3. Ax
4. $A^T A$
5. $Tr(A^T A)$

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 statement is false,please give the correct form.

1. $x(y + z)^T = z^T + y^T x$
2. $x^T x = xx^T$
3. $x^T Ay = y^T A^T x$
4. $x^T y^T z = (xy)^T z$
5. $AB = BA$
6. $(AB)C = A(BC)$
7. $A(B + C)^T = B^T A + C^T A$
8. $(AB)^T = A^T B^T$

2 Gradients and Hessians of Linear and Quadratic Functions

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

1. $f(x) = a^T x + \alpha$ (linear)
2. $f(x) = a^T x + a^T A x + x^T A^T b$ (more linear forms)
3. $f(x) = x^T x + x^T W x + x^T A B x$ (quadratic forms)
4. $f(x) = \frac{1}{2}(Ax - b)^T W (Ax - b)$ (weighted least squares)
5. $f(x) = \frac{\lambda}{2} \|x\|^2 + \frac{1}{2} \sum_{i=1}^{n} (a_i^T x - b_i)^2$ (L2-regularized least squares)