CS229
Problem Set 0
Linear Algebra and Multivariable Calculus

Kangwei Ling @kevinkwl

February 2, 2017

1. Gradient and Hessians

   Solution.

   (a) Note that $x^T Ax = \Sigma_{1 \leqslant i,j \leqslant n} A_{ij} x_i x_j$, and since A is a symmetric matrix, $x^T Ax = \Sigma_{i=1}^{n} A_{ii} x_i^2 + \Sigma_{i=1}^{n-1} \Sigma_{j=i+1}^{n} 2A_{ij} x_i x_j$. Thus,

   $$\nabla f(x) = \begin{bmatrix} \vdots \\ 1/2(2A_{ii}x_i + \Sigma_{1 \leqslant j \leqslant n, j \neq i} 2A_{ij} x_j) + b_i \\ \vdots \end{bmatrix} = \begin{bmatrix} \vdots \\ \Sigma_{1 \leqslant j \leqslant n} A_{ij} x_j + b_i \\ \vdots \end{bmatrix} = Ax + b$$

   (b) Use the **Chain Rule**.

   $$\nabla f(x) = \begin{bmatrix} g'(h(x)) \cdot \frac{\partial}{\partial x_1} h(x) \\ \vdots \end{bmatrix}$$

   (c) from the answer in (a), it's easy to find out that,

   $$\nabla^2 f(x) = A$$

   (d)

   $$\nabla f(x) = g'(a^T x)a$$
   $$\nabla^2 f(x) = g''(a^T x)aa^T$$

   $\square$

2. Positive definite matrices

   (a) Firstly, $A^T = (zz^T)^T = zz^T = A$. Then for all $x \in \mathbb{R}^n$, $x^T Ax = x^T zz^T x = (z^T x)^2 \geqslant 0$.

(b) A is PSD. Consider the linear system $Ax = 0$, notice that for all x', s.t., $Ax' = 0$, then $zz^Tx' = 0$. Note that z^Tx' is a scalar, and because z is non-zero, $z^Tx' = 0$.

Therefore, the null space of A is $\{x \in \mathbb{R}^n \mid z^Tx = 0\}$. As the null space is a hyperplane perpendicular to non-zero vector z, the dimension of A is $n-1$. Thus $\text{rank}(A) = n-(n-1) = 1$

(c) $BAB^T \in \mathbb{R}^{m\times m}$, $(BAB^T)^T = BA^TB^T = BAB^T$. For all $x \in \mathbb{R}^m, B^Tx \in \mathbb{R}^n$, and $x^TB = (B^Tx)^T$. Let $y = B^Tx, y \in \mathbb{R}^n$, then for all such $x, y^TAy \geqslant 0$, that is, $x^TBAB^Tx \geqslant 0$. Therefore, BAB^T is PSD.

3. Eigenvectors, eigenvalues, nd the spectral theorem

(a) Let $I^{(i)}$ be the i^{th} column of I. Note that $T^{-1}T = I, T^{-1}t^{(i)} = I^{(i)}$. That is,

$$T^{-1}t^{(i)} = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$

and it's also obvious that

$$\frac{1}{\lambda_i}\Lambda = \begin{bmatrix} \frac{\lambda_1}{\lambda_i} & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & \frac{\lambda_n}{\lambda_i} \end{bmatrix}$$

therefore,

$$\frac{1}{\lambda_i}\Lambda T^{-1}t^{(i)} = T^{-1}t^{(i)}$$

i.e. $At^{(i)} = \lambda_i t^{(i)}$

(b) Since $U^TU = I$, $U^{-1} = U^T$, the statement follows immediately from the result of (a).

(c) A is PSD, then $A = U\Lambda U^T$ by the **spectral theorem**.For each $u^{(i)}$, we know that $Au^{(i)} = \lambda_i u^{(i)}$. And

$$(t^{(i)})^TAt^{(i)} \geqslant 0$$

since

$$(t^{(i)})^TAt^{(i)} = \lambda_i(t^{(i)})^Tt^{(i)} = \lambda_i|t^{(i)}|^2$$

$\lambda_i \geqslant 0$