PS0 Solution (Harold Mei, hmei0411@stanford.edu)

1 Gradient and Hessians

(a) $\nabla f(x) = Ax + b$

(b) $\nabla f(x) = g'(h(x))\nabla h(x)$

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

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

2 Positive definite matrices

(a) Show that $A = zz^T$ is positive semidefinite.

PROOF: suppose vector x $\neq$ 0; the qustion is to just prove $x^T Ax \geq 0$.

$$x^T Ax = x^T zz^T x \tag{1}$$

Since $x^T z = z^T x = c$ is just a scalor, $x^T Ax = c^2 \geq 0$.

(b) Let $A = zz^T$ and z is non-zero, nullspace and rank of A.

Nullspace of A is just the solution space for $Ax = 0$.
Since $x^T Ax = x^T zz^T x = (x^T z)^2 \geq 0$, x can only be $\vec{0}$ if $Ax = 0$, so nullspace of A is $\{\vec{0}\}$.

A is full rank. Sicne A is positive definite, it is a full rank matrix.
Otherwise some column can be written as a linear combination of all other columns and hence an all 0 column can be generated. Suppose after some linear transformation the ith column is all 0, pick a vector x to have its ith component non-zero while all others are zeros, $x^T Ax = 0$ which contradicts the PD definition.

(c) Prove that BAB^T is PSD.

Given any vector $x \in \mathbb{R}^{1\times m}$, we have $y = xB \in \mathbb{R}^{1\times n}$

$$xBAB^T x = yAy^T \geq 0 \tag{2}$$

Q.E.D.

3 Eigenvectors, eigenvalues and the spectral theorem

(a) show that $At^{(i)} = \lambda_i t^{(i)}$ given diagnalizable matrix $A = T\Lambda T^{-1}$

Since $AT = T\Lambda$ and $T = [t^{(1)}, t^{(2)}, \ldots, t^{(n)}]$, which is just:

$$A[t^{(1)}, t^{(2)}, \ldots, t^{(n)}] = [t^{(1)}, t^{(2)}, \ldots, t^{(n)}]\Lambda \tag{3}$$

$$\Rightarrow [At^{(1)}, At^{(2)}, \ldots, At^{(n)}] = [\lambda_1 t^{(1)}, \lambda_2 t^{(2)}, \ldots, \lambda_n t^{(n)}] \tag{4}$$

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

From (3) to (4) the multiplication applies on the ith component of every row vector for λ_i, which is just the ith column vector $t^{(i)}$.

(b) Properties of symmetric matrices and orthogonal matrices

PROOF:
By the spetral theorem, given a symmetric matrix A, it can be written as $A = U\Lambda U^T$ for an orthogonal matrix U, $U^TU = I$
Orthogonal matrix has a nice property $U^T = U^{-1}$, which can be pluged into $A = U\Lambda U^T$ and change the form to the following:

$$AU = U\Lambda$$

This is just the same as in 3.(a) which gives the result:

$$Au^{(i)} = \lambda_i u^{(i)}$$

where $\Lambda = diag(\lambda_1, \lambda_2, \ldots, \lambda_n)$ and ($u^{(i)}, \lambda_i$) are eigenvector/eigenvalue pairs.

(c) Show that if A is PSD then $\lambda_i(A) \geq 0$ for each i

Suppose A is PSD, then $\forall x, x^TAx \geq 0$;
For any of its eigenvector/eigenvalue pair ($u^{(i)}, \lambda_i$):

$$Au^{(i)} = \lambda_i u^{(i)}$$

$$\Rightarrow {u^{(i)}}^T Au^{(i)} = {u^{(i)}}^T \lambda_i u^{(i)} = \lambda_i \|u^{(i)}\|_2^2 \geq 0$$

$$\Rightarrow \lambda_i(A) \geq 0$$

Q.E.D.