Math 650, Foundations of Optimization, Spring 2010

Sandipan Dey,
Homework Assignment - 2

March 4, 2010

Problem 1 Solution

$f(x,y,z) = xyz.e^{-x-y-z}$

By the first order necessary condition for critical points, we have

$\nabla f(x,y,z) = 0 \Rightarrow$

$$
\begin{aligned}
f_x(x,y,z) &= \frac{\partial}{\partial x} f(x,y,z) = yz.(1-x).e^{-x-y-z} = 0 \Rightarrow y = 0 \vee z = 0 \vee x = 1 \\
f_y(x,y,z) &= \frac{\partial}{\partial y} f(x,y,z) = zx.(1-y).e^{-x-y-z} = 0 \Rightarrow z = 0 \vee x = 0 \vee y = 1 \\
f_z(x,y,z) &= \frac{\partial}{\partial z} f(x,y,z) = xy.(1-z).e^{-x-y-z} = 0 \Rightarrow x = 0 \vee y = 0 \vee z = 1
\end{aligned}
$$

Hence, the set of all possible critical points:
$\{(0,0,0),(1,0,0),(0,0,1),(0,1,0),(1,1,1),(a,0,0),(0,a,0),(0,0,a)\}$, when $0 \neq a \in \Re$, i.e., $\{(0,0,0),(1,1,1),(a,0,0),(0,a,0),(0,0,a)\}$, when $0 \neq a \in \Re$

Now let's examine the second order sufficient conditions for finding the type of the critical point. For this we need to compute the Hessian matrix of $f(x,y,z)$

$$
H\left(f\left(x,y,z\right)\right) = e^{-x-y-z} \begin{pmatrix} (2-x)yz & (1-x)(1-y)z & (1-z)(1-x)y \\ (1-x)(1-y)z & (2-y)zx & (1-y)(1-z)x \\ (1-z)(1-x)y & (1-y)(1-z)x & (2-z)xy \end{pmatrix}
$$

Hence, we have,

$H\left(f\left(0,0,0\right)\right)=\begin{pmatrix}0&0&0\\0&0&0\\0&0&0\end{pmatrix}$, not a strictly positive/negative definite matrix, or both a positive/negative semidefinite matrix.

$H\left(f\left(0,0,1\right)\right)=\begin{pmatrix}0&\frac{1}{e}&0\\\frac{1}{e}&0&0\\0&0&0\end{pmatrix}$, $|H-\lambda I|=0$ gives the characteristic polynomial $-\lambda^3+\frac{\lambda}{e^2}=0 \Rightarrow \lambda=0,\pm\left(\frac{1}{e}\right) \Rightarrow$ the Hessian matrix has both positive and negative eigenvalues $\Rightarrow$ an indefinite matrix $\Rightarrow (0,0,1)$ is a saddle point by the second order sufficient condition. (Also, by Sylvester's theorem we can see that $det(H_1)=0, det(H_2)=-\frac{1}{e^2}$, hence not all leading principle diagonal matrices have strictly positive determinant).

Similarly, at points $(0,1,0)$ and $(1,0,0)$, the Hessian matrix has both $+ve$ and $-ve$ eigenvalues $\Rightarrow$ the Hessian matrix is an indefinite matrix $\Rightarrow (0,1,0)$ and $(1,0,0)$ are saddle points.

Applying similar logic, we have all the critical points of the form $(0,0,a)$, $(0,a,0)$, $(a,0,0),\ a\neq 0$ as saddle points.

$H\left(f\left(0,0,1\right)\right)=\begin{pmatrix}e^3&0&0\\0&e^3&0\\0&0&e^3\end{pmatrix}$, $|H-\lambda I|=0$ gives the characteristic polynomial $(e^3-\lambda)^3=0 \Rightarrow \lambda=e^3,e^3,e^3 \Rightarrow$ all the eigenvalues are strictly $+ve \Rightarrow$ H is a positive definite matrix $\Rightarrow (1,1,1)$ is a minimum point. (Also, by Sylvester's theorem we can see that $det(H_i)=e^{3i}>0,\ i=1,2,3$, hence all leading principle diagonal matrices have strictly positive determinant).

$(1,1,1)$: local minimum (no other local minima $\Rightarrow$ a global minimum too).
$(0,0,a),(0,a,0),(a,0,0)$: saddle points when $a\neq 0$.
$(0,0,0)$: can't say anything.

Problem 2 Solution

Given $f(x)=\frac{1}{2}\langle Ax,x\rangle+\langle c,x\rangle+a=(x^T\ 1)\begin{pmatrix}A&c\\c^T&a\end{pmatrix}\begin{pmatrix}x\\1\end{pmatrix}$, with $A_{n\times n}$ symmetric $\Rightarrow \nabla f(x)=\frac{\partial f}{\partial x}=x^TA+c^T$, $H(f(x))=\nabla^2 f(x)=A$.

Also, $f(x)$ is bounded from below $\Rightarrow \exists l_b\in\Re^n \mid f(x)\geq l_b,\ \forall x\in\Re^n$.
$\Rightarrow f(0)=a\geq l_b \Rightarrow a-l_b\geq 0$.

Also, A is the symmetric $\Rightarrow A$ is diagonalizable $\Rightarrow A = u\Lambda u^T$, with diagonal Λ and $u = [u_1 u_2 \dots u_n]$.

$\Rightarrow f(x) = \frac{1}{2}\left\langle u\Lambda u^T x, x\right\rangle + \langle c, x\rangle + a = \frac{1}{2}\left\langle \Lambda u^T x, u^T x\right\rangle + \langle c, x\rangle + a$
$= \frac{1}{2}\langle \Lambda y, y\rangle + \langle c, x\rangle + a$, where $y = u^T x,\ u^T : \Re^n \rightarrow \Re^n$.

$\Rightarrow f(x) = \frac{1}{2}\sum_i \lambda_i y_i^2 + \sum_i c_i x_i + a \geq l_b,\ \forall x \in \Re^n$.

A is p.s.d.: Proof by contradiction

Let's assume to the contrary, i.e., A is not positive semidefinite $\Rightarrow \exists k \mid \lambda_k < 0$ and λ_k is an eigenvalue of A.

Now let's examine $f(x)$ at the point $x^* = (0, 0, \dots, z, \dots, 0, 0)$, where $(x_i = z$, if $i = k)$ and $(x_i = 0,\ \forall i \neq k)$, where $z \in \Re$.
$\Rightarrow f(x^*) = \frac{1}{2}\lambda_k \sum_i u_{ik}^2 z^2 + c_k z + a$. (since $x_k = z$)
Let's denote $p = \sum_i u_{ik}^2 \geq 0$ and $q = c_k$, and consider the roots of the quadratic

$f(x^*) - l_b = \frac{1}{2}p\lambda_k z^2 + qz + a - l_b = 0 \Rightarrow (z_1, z_2) = \frac{-q \pm \sqrt{q^2 - 2p\lambda_k(a - l_b)}}{p\lambda_k}$.

We notice that $p, q^2, a - l_b \geq 0 \wedge \lambda_k < 0 \Rightarrow z_1, z_2 \in \Re$ (discriminant ≥ 0).

If we choose z (we can always choose any $z \in \Re$) s.t. $z \in (z_1, z_2)$, i.e., $\frac{-q - \sqrt{q^2 - 2p\lambda_k(a - l_b)}}{p\lambda_k} = z_1 < z < z_2 = \frac{-q + \sqrt{q^2 - 2p\lambda_k(a - l_b)}}{p\lambda_k}$

$\Rightarrow f(x^*) - l_b = (z - z_1)(z - z_2) < 0 \Rightarrow f(x^*) < l_b$, a contradiction.

Hence, our initial assumption was wrong, i.e., the symmetric matrix A can not have any negative eigenvalue $\Rightarrow$ all the eigenvalues of A must be non-negative real numbers $\Rightarrow A$ is positive semi-definite (Proved).

f has global minimum in $\Re^n$

We consider the compact sublevel subset $\{f(x) \in \Re^n : f(x) \leq l_b\}$ of $\Re^n$ and see that $f : \Re^n \rightarrow \Re$ is a continuous (convex) function defined on the metric space $\Re^n$. Hence, f achieves a global minimizer on $\Re^n$.

Problem 3 Solution

Part (a)

Given $p(x) = \langle Bx, x\rangle + 2\langle b, x\rangle + c = (x^T \;\; 1)\begin{pmatrix} B & b \\ b^T & c \end{pmatrix}\begin{pmatrix} x \\ 1 \end{pmatrix}$, with $B_{n\times n}$ is positive definite.

By the first order necessary condition for the critical point at x^*, we have,
$\nabla p(x^*) = \left[\frac{\partial p}{\partial x}\right]_{x^*} = 2({x^*}^T B + b^T) = 0 \Rightarrow {x^*}^T BB^{-1} = -b^T B^{-1}$
$\Rightarrow x^* = -B^{-1}b$
$\Rightarrow p(x^*) = \left\langle b, B^{-1}b\right\rangle + 2\left\langle b, B^{-1}b\right\rangle + c = c - \left\langle B^{-1}b, b\right\rangle,$
($\exists B^{-1}$, since B is positive definite, i.e, all eigenvalues striclty positive, hence non-singular and invertible, also B is symmetric, $B^T = B$).

Also, $p(x)$ is strictly positive in $R^n \Rightarrow p(x) > p(x^*) = c - \left\langle B^{-1}b, b\right\rangle > 0$, $\forall x \neq x^* \in \Re^n$.

Now, the Hessian $H(p(x)) = \nabla^2 p(x) = 2B$ is positive definite, hence by second order sufficient condition, x^* is a local minimum.

As we can see from the first order necessary condition x^* is the only critical point for the function $p(x)$ and by the second order sufficient condition we see that it's the minimum point, hence the function is having the unique global mimimum at x^*.

Part (b)

Choose $d = x^* = -B^{-1}b \in \Re^n$.

$$\begin{pmatrix} I_n & 0 \\ d^T & 1 \end{pmatrix}.A.\begin{pmatrix} I_n & d \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} I_n & 0 \\ -b^T B^{-1} & 1 \end{pmatrix}\begin{pmatrix} B & b \\ b^T & c \end{pmatrix}\begin{pmatrix} I_n & -B^{-1}b \\ 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} B & b \\ 0 & -b^T B^{-1}b + c \end{pmatrix}\begin{pmatrix} I_n & -B^{-1}b \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} B & 0 \\ 0 & -b^T B^{-1}b + c \end{pmatrix}$$

$$\Rightarrow det\left(\begin{pmatrix} I_n & 0 \\ d^T & 1 \end{pmatrix}.A.\begin{pmatrix} I_n & d \\ 0 & 1 \end{pmatrix}\right) = det\begin{pmatrix} B & 0 \\ 0 & -b^T B^{-1}b + c \end{pmatrix}$$

$$\Rightarrow det\begin{pmatrix} I_n & 0 \\ d^T & 1 \end{pmatrix}.det(A).det\begin{pmatrix} I_n & d \\ 0 & 1 \end{pmatrix} = det\begin{pmatrix} B & 0 \\ 0 & -b^T B^{-1}b + c \end{pmatrix}$$
$\Rightarrow det(A) = det(B).(c - b^T B^{-1}b)$ (Proved)

Part (c)

Proof (⇔) by induction on the dimension n of matrix $A_{n\times n}$

Base case: for $n = 1$ the proof is trivial (since A contains only one element).

Induction Hypothesis: Let's assume the result is true for n.

Induction Step: Let's prove the result for $n + 1$.
Let's consider the $(n + 1) \times (n + 1)$ symmetric matrices (A) of the form

$A = \begin{pmatrix} B & b \\ b^T & c \end{pmatrix}$, where B is a poditive definite matrix as in part (b). We notice that all the positive $(n+1)\times(n+1)$ definite matrices will be of this form (hence will be a subset of this set) and they must be symmetric.

Now, consider the quadratic form $\langle Ay, y\rangle$ of the matrix A, where $y \in \Re^{n+1}$. We notice that any such $y \in \Re^{n+1}$ can be represented as $[x \ \ a]$, where $a \in \Re$.

1. When $a = 0$, $\langle Ay, y\rangle = \langle Bx, x\rangle \wedge B$ is positive definite $\Leftrightarrow A$ is positive definite (all leading principle minors are positive by induction hypothesis).

2. When $a \neq 0$, $\langle Ay, y\rangle = y^T Ay = (x^T \ 1) \begin{pmatrix} B & b \\ b^T & c \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix} = p(x)$

   From part (b), $< Ay, y >= p(x) > 0 \Leftrightarrow c - \left\langle B^{-1}b, b\right\rangle > 0$.
   Also, B is positive definite $\Rightarrow det(B) > 0$ (by induction hypothesis)
   $\Rightarrow detA = detB.(c - \left\langle B^{-1}b, b\right\rangle) > 0$.Also, all leading principle minors of B are positive by induction hypothesis and $det(A) > 0 \Leftrightarrow$ all leading principle minors of A are positive $\Leftrightarrow A$ is postive definite. (Proved)