CPSC 340 Assignment 0

Kyle Leeners, 30591119, w0u0b

January 19, 2018

1. Assume $||\cdot||_{\frac{1}{2}}$ is a norm. Then it must be the case that $||u+v||_{\frac{1}{2}} \leq ||u||_{\frac{1}{2}} + ||v||_{\frac{1}{2}}$. Let $u = [1,0]^T, v = [0,1]^T$.

$$||u+v||_{\frac{1}{2}} = (2 \cdot \sqrt{1})^2 = 4$$
$$||u||_{\frac{1}{2}} = ||v||_{\frac{1}{2}} = 1$$
$$||u+v||_{\frac{1}{2}} = 4 > ||u||_{\frac{1}{2}} + ||v||_{\frac{1}{2}} = 2$$

So our assumption was incorrect and $||\cdot||_{\frac{1}{2}}$ cannot be a norm

2. (a) $||A|| = \sqrt{\lambda_{max}(A^TA)} = \sqrt{\lambda_{max}(AA^T)} = ||A^T||$

(b) $||A||_F = \sum_{i=1}^n \sum_{j=1}^n (a_{i,j})^2$. This is the sum of squares of every value of the matrix A which corresponds $Tr(A^TA) = \sum_{i=1}^n \lambda_i(A^TA)$

3. Let $g(v) = ||x||^2 - 1$. So we wish to maximize $f(x) = x^tAx$ subject to $g(v) = 0$

$$\nabla f = \lambda \nabla g$$

$$\nabla g = 2x$$

$$\nabla f = \nabla x^TAx = \frac{\partial(x^tAx)}{\partial x} + \frac{\partial(Ax)^T}{\partial x} \cdot \frac{\partial(x^TAx)}{\partial y} = Ax + \frac{\partial(x^TA^T)}{\partial x} \cdot \frac{\partial(y^Tx)}{\partial y} = Ax + A^Tx = 2Ax$$
$$\Rightarrow 2Ax = \lambda 2x \Rightarrow Ax = \lambda x$$

This implies that x must be an eigenvector of A.

$$f(x) = x^TAx = x^T\lambda x = \lambda x^Tx = \lambda ||x||^2 = \lambda$$

Therefore, the max of $f(x)$ subject to $||x||^2 = 1$ is equivalent to λ_{max}

4. (a) Let λ, x be anyt eigenvector, eigenvalue pair of A, then:

$$Ax = \lambda x$$

$$\Rightarrow x^TAx = \lambda x^Tx$$
$$\Rightarrow \lambda = \frac{x^TAx}{||x||^2} = \frac{x^TBB^Tx}{||x||^2} = \frac{(B^Tx)^TB^Tx}{||x|^2} = \frac{||Bx||^2}{||x||^2} \geq 0$$

(b) First we will prove if A is positive definite then B will have full row rank. Since A is positive definite all eigen values are positive. This implies that the matrix A will have full row/column rank (since it is symmetric). Since $A = BB^T$ and $B \in \mathbb{R}^{n \times k}$ B must have full row-rank.

Now proof if B has full row rank, A will be positive definite. We know $rank(B) = rank(B^T)$. Furthermore, we know $rank(A) = rank(BB^T) = min\{rank(B), rank(B^T)\} = rank(B)$. Since $B \in \mathbb{R}^{n \times k}$ and is full row rank, $rank(B) = n$ which implies $rank(A) = n$. Since $A \in \mathbb{R}^{n \times n}$ A must be full rank and have a nullity of 0. This implies all eigenvalues are greater than 0 so A is positive definite

5. We wish to show $D \succeq 0 \leftrightarrow c - b^T A^{-1} b \geq 0$

   If $D \succeq 0 \rightarrow \forall_{z \in \mathbb{R}^{n+1}}, z^T D z \geq 0$
   Let $x \in \mathbb{R}^n, y \in \mathbb{R}, z = (x, y)^T$
   This implies that $z^T D z = f(x, y) = x^T A x + 2yb^t x + cy^2$
   Consider two possibilities

   (a) y $= 0 \rightarrow f(x, 0) = x^t A x > 0$ since $A \succ 0 \rightarrow D \succ 0$

   (b) $y \neq 0$ then after computing the gradient and Hessian we obtain:
   $\nabla f(x, y) = 2Ax + 2yb$ and $\nabla^2 f(x, y) = 2A \succ 0$ since $A \succ 0$
   This implies that the stationary point must be a global minimum and the minimizer will be:
   $\nabla f(x, y) = 2Ax + 2yb = 0 \rightarrow Ax = -yb \rightarrow x^* = -yA^{-1}b$

   Evaluating at the minimum gives:
   $f(x^*, y) = (-yA^{-1}b)^T A(-yA^{-1}b) + 2yb^T(-yA^{-1}b) + cy^2$
   $= y^2(b^T A^{-1} A A^{-1} b) - 2y^2(b^T A^{-1} b) + y^2 c$
   $= y^2(c - b^T A^{-1} b)$

   So if $f(x, y) \geq 0 \rightarrow c - b^T A^{-1} b \geq 0$

   Assume $(c - b^T A^{-1} b) \geq 0$. From the previous proof, we know $y^2(c - b^T A^{-1} b)$ is the global minimum x^* for $f(x, y) = x^T A x + 2yb^t x + cy^2$. This implies that $\forall_{x,y}, f(x, y) \geq 0$. Recall that $f(x, y) = (x, y)D(x, y)^T = z^T D z$. Since we know $z^T D z \geq 0$ it follows that $D \succeq 0$.

6. Note when I say "evaluate the Hessian at the roots", I mean find the eigenvalues of the Hessian after inputing the stationary points

   (a) $\frac{\partial f(x_1,x_2)}{\partial x_1} = 16x_1(4x_1^2 - x_2)$
   $\frac{\partial f(x_1,x_2)}{\partial x_2} = -2(4x_1^2 - x_2)$
   Roots : $(x_1, 4x_1)$
   Since $f(x_1, x_2)$ is squared we know these roots will correspond to the non-strict global minimum which in this case is 0

   (b) $\frac{\partial f(x_1,x_2,x_3)}{\partial x_1} = 4x_1^3 - 4x_1$
   $\frac{\partial f(x_1,x_2,x_3)}{\partial x_2} = 2x_2 + 2x_3$
   $\frac{\partial f(x_1,x_2,x_3)}{\partial x_3} = 2x_2 + 4x_3$
   Roots: $(0, 0, 0), (1, 0, 0), (-1, 0, 0)$
   $\frac{\partial^2 f(x_1,x_2,x_3)}{\partial x_1^2} = 12x_1^2 - 4$
   $\frac{\partial^2 f(x_1,x_2,x_3)}{\partial x_2^2} = 2$
   $\frac{\partial^2 f(x_1,x_2,x_3)}{\partial x_3^2} = 4$
   $\frac{\partial^2 f(x_1,x_2,x_3)}{\partial x_1 x_2} = 0$
   $\frac{\partial^2 f(x_1,x_2,x_3)}{\partial x_1 x_3} = 0$
   $\frac{\partial^2 f(x_1,x_2,x_3)}{\partial x_2 x_3} = 2$
   Evaluating the Hessian at the roots gives the following:

Saddle point: $(0,0,0)$ Non-strict global minimizers: $(1,0,0),(-1,0,0)$

(c) $\frac{\partial f(x_1,x_2)}{\partial x_1} = 6x_1x_2$
$\frac{\partial f(x_1,x_2)}{\partial x_2} = 6x_2^2 - 12x_2 + 3x_1^2$
Roots: $(0,0),(0,2)$
$\frac{\partial^2 f(x_1,x_2)}{\partial x_1^2} = 6x_2$
$\frac{\partial^2 f(x_1,x_2)}{\partial x_2^2} = 12x_2 - 12$
$\frac{\partial^2 f(x_1,x_2)}{\partial x_1 x_2} = 6x_1$
Evaluating the Hessian at the roots gives the following:
Strict local minimizer: $(0,2)$ Non-strict local maximizer: $(0,0)$

(d) $\frac{\partial f(x_1,x_2)}{\partial x_1} = 4x_1^3 + 4x_1x_2 - 8x_1 - 8$
$\frac{\partial f(x_1,x_2)}{\partial x_2} = 2x_1^2 + 2x_2 - 8$
Roots: $(1,3)$
$\frac{\partial^2 f(x_1,x_2)}{\partial x_1^2} = 12x_1^2 + 4x_2 - 8$
$\frac{\partial^2 f(x_1,x_2)}{\partial x_2^2} = 2$
$\frac{\partial^2 f(x_1,x_2)}{\partial x_1 x_2} = 4x_1$
Evaluating the Hessian at the roots gives the following:
Strict global minimizer: $(1,3)$

(e) $\frac{\partial f(x_1,x_2)}{\partial x_1} = 4(x_1 - 2x_2)^3 + 64x_2$
$\frac{\partial f(x_1,x_2)}{\partial x_2} = -8(x_1 - 2x_2)^3 + 64x_1$
Roots: $(0,0),(1,-\frac{1}{2}),(-1,\frac{1}{2})$
$\frac{\partial^2 f(x_1,x_2)}{\partial x_1^2} = 12(x_1 - 2x_2)^2$
$\frac{\partial^2 f(x_1,x_2)}{\partial x_2^2} = 48(x_1 - 2x_2)$
$\frac{\partial^2 f(x_1,x_2)}{\partial x_1 x_2} = 64 - 24(x_1 - 2x_2)^2$
Evaluating the Hessian at the roots gives the following:
Saddle point: $(0,0)$ Non-strict global minimizer: $(1,-\frac{1}{2}),(-1,\frac{1}{2})$

(f) $\frac{\partial f(x_1,x_2)}{\partial x_1} = 4x_1 - 2x_2 + 2$
$\frac{\partial f(x_1,x_2)}{\partial x_2} = 6x_2 - 2x_1 - 3$
Roots: $(-\frac{3}{10},\frac{2}{5})$
$\frac{\partial^2 f(x_1,x_2)}{\partial x_1^2} = 4$
$\frac{\partial^2 f(x_1,x_2)}{\partial x_2^2} = 6$
$\frac{\partial^2 f(x_1,x_2)}{\partial x_1 x_2} = -2$
Evaluating the Hessian at the roots gives the following:
Strict Global minimizer: $(-\frac{3}{10},\frac{2}{5})$

(g) $\frac{\partial f(x_1,x_2)}{\partial x_1} = 2x_1 + 4x_2 + 1$
$\frac{\partial f(x_1,x_2)}{\partial x_2} = 4x_1 + 2x_2 - 1$
Roots: $(\frac{1}{2},-\frac{1}{2})$
$\frac{\partial^2 f(x_1,x_2)}{\partial x_1^2} = 2$
$\frac{\partial^2 f(x_1,x_2)}{\partial x_2^2} = 2$

$\frac{\partial^2 f(x_1,x_2)}{\partial x_1 x_2} = 4$
Evaluating the Hessian at the roots gives the following:
Saddle point at: $(\frac{1}{2}, -\frac{1}{2})$

7. Assume $b \in Range(A)$, we wish to prove that $f(x)$ is bounded below

   $f(x) = x^T Ax + 2b^T x + c \to \nabla f(x) = 2Ax + 2bx$. Assume $b \in Range(A) \to \exists_y, s.t, Ay = b$.
   Let $x = -y \to \nabla f(-y) = -2Ay + 2b = -2b + 2b = 0$. This implies that at $x = -y$ we have a stationary point. Now examining the Hessian we find $\nabla^2 f(x) = 2A \succeq 0$ since $A \succeq 0$. This implies that our stationary point is a global minimum and $f(x)$ is bounded below.

   Assume $f(x)$ is bounded below, we need to show $b \in Range(A)$.

   We know from the previous proof that since $A \succeq 0$, $\nabla^2 f(x) = 2A \succeq 0$. This means there exists some global minimizer x^* where $\nabla f(x^*) = 2Ax^* + 2b \to Ax^* = -b \to b \in Range(A)$