Jesse Cai
MATH 164 Homework 1
304634445

1. **6.1 Deterime if the given point is definitely a local minimizer, definitely not, or possibly**
   (a) Definitely not a local minimizer, violates FONC $d = [1, 1]$ is feasible and $d^T \nabla f(x^*) = \sqrt{2} > 0$.
   (b) Possibly a local minimizer, x^* satisfies FONC.
   (c) Definitely a local minimizer, this satisfies SOSC, as $x*$ is an interior point, $\nabla f(x^*) = 0$ and $F(x^*) = I([1, 2]^T) = 3 > 0$.
   (d) Definitely not a local minimizer, violates SONC $d = [0, 1]^T$ is feasible and $d^T F(x^*)d = -1 \leq 0$.

2. **6.2 Show that if x^* is a global minimizer over Ω and $x^* \in \Omega' \subset \Omega$ then x^* is also a global minimizer over Ω'**
   From Def 6.1, $\forall x \in \Omega, f(x) \geq f(x^*)$. Pick some element in Ω'. $y \in \Omega' \implies y \in \Omega \implies f(y) \geq f(x^*)$. So $\forall y \in \Omega', f(y) \geq: (x^*) \implies x^*$ is a global minimizer over Ω'.

3. **6.4 Show** $x_0 + \arg\min_{x\in\Omega} f(x) = \arg\min_{y\in\Omega'} f(y)$
   Let $x^* = \arg\min_{x\in\Omega} f(x), y^* = \arg\min_{y\in\Omega'} f(y)$. Suppose $x^* \neq y^* + x_0$. Then either $x^* \leq y^* + x_0$ or $y^* + x_0 \leq x^*$.
   WLOG consider $x^* \leq y^* + x_0 \implies x^* - x_0 \leq y^*, x^* - x_0 \in \Omega'$, which is a contradiction, as we defined y^* as the mininum.

4. **6.8 Consider** $f(x) = x^T \begin{bmatrix} 1 & 2 \\ 4 & 7 \end{bmatrix} x + x^T [35] + 6$
   (a) $\nabla f(x) = \left(\begin{bmatrix} 1 & 2 \\ 4 & 7 \end{bmatrix} + \begin{bmatrix} 1 & 4 \\ 2 & 7 \end{bmatrix} \right) \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 3 \\ 5 \end{bmatrix} = \begin{bmatrix} 11 \\ 25 \end{bmatrix}$
   $F(x) = \begin{bmatrix} 2 & 6 \\ 6 & 14 \end{bmatrix}$
   (b) The unit vector in the direction of max increase is given by normalizing $\nabla f(x)$. $D_d f(x) = \begin{bmatrix} \frac{11}{\sqrt{746}} \\ \frac{25}{\sqrt{746}} \end{bmatrix} \nabla f(x) = \sqrt{746}$
   (c) Solving $\nabla f(x) = 0$ yields $x = \begin{bmatrix} 2 & 6 \\ 6 & 14 \end{bmatrix}^{-1} \begin{bmatrix} -3 \\ -5 \end{bmatrix} = \begin{bmatrix} \frac{3}{2} \\ -1 \end{bmatrix}$
   However, this point does not satisfy SONC as for $d = -x, d^T F(x)d = -0.5 < 0$.

5. **6.11 Consider minimize $-x_2^2$ s.t.** $|x_2| \leq x_1^2, x_1 \geq 0$
   (a) $\nabla f(x) = \begin{bmatrix} 0 \\ -2x_2 \end{bmatrix}, \nabla f(0) = 0$ so for any feasible direction d $d^T \nabla f(0) = 0$, so FONC is satisfied.
   (b) 0 is a local maxamizer but not a strict local maxamizer, as we can move along the x_1 axis for more maxima.

6. **6.12 Consider minimizing** $f(x) = 5x_2$ **s.t.** $x_1^2 + x_2 \geq 1$
   (a) $\nabla f([0, 1]^T) = [0, 5]^T$. The set of feasible directions is $\{[d_1, d_2] : d_1, d_2 > 0\}$. But then $d^T \nabla f([0, 1]^T) = 5d_2 > 0$ because $d_2 > 0$. So the FONC is satisfied.
   (b) Note that $F([0, 1]^T) = 0$ so $\forall d : d^T F(x)d = 0$ and SONC is satisfied.
   (c) No $x^* = [0, 1]^T$ is not a local

7. **6.16 Consider minimizing** $f(x) = 4x_1^2 - x_2^2$ **s.t.** $x_1^2 - 2x_1 - x_2 \geq 0, x_1 \geq 0, x_2 \geq 0$
   (a) $\nabla f(x) = [8x_1, -2x_2]^T$ so evaluated at $x^* = [0, 0]^T, \nabla f(x^*) = [0, 0]^T$ and as such $\forall d : d^T \nabla f(x^*) = 0$ and FONC is satisfied.
   (b) Note that $F = \begin{bmatrix} 8 & 0 \\ 0 & -2 \end{bmatrix}$ and the set of feasible directions is $\{[d_1, d_2]^T : d_2 \leq 2d_1, d_1 \geq 0, d_2 \geq 0 \}$
   Then we know that $d^T Fd = 8d_1^2 - 2d_2^2 \geq 8d_1^2 - 2(2d_1^2) = 0$ so SONC is satisfied.

8. **8.1 Perform two iterations of the minimization of $f(x_1, x_2) = x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_1^2 + x_2^2 + 3$**

   We'll rewrite this in standard form $f(x) = x^T \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix} x + x^T[1, \frac{1}{2}] + 3$

   Then we know that $x_{n+1} = x_n - \alpha_n \nabla f(x_n)$ and $\alpha_n = \frac{g_n^T g_n}{g_n^T F g_n}$.

   $g_0 = \nabla f(x_0) = [1, \frac{1}{2}]^T$ so $x_1 = -\frac{5}{6}[1, \frac{1}{2}]^T = [-\frac{5}{6}, -\frac{5}{12}]^T$

   $g_1 = \nabla f(x_1) = [\frac{1}{6}, -\frac{1}{3}]^T$ so $x_2 = [-\frac{5}{6}, -\frac{5}{12}]^T - \frac{5}{9}[\frac{1}{6}, -\frac{1}{3}]^T = \begin{bmatrix} -\frac{25}{27} \\ -\frac{25}{108} \end{bmatrix}$

   Analytically the solution is found when $x^* = F^{-1}(-b) = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} -1 \\ -\frac{1}{2} \end{bmatrix} = [-1, -\frac{1}{4}]^T$

9. **8.8 Find the largest range of values for α for which the algorithm is globally convergent.**

   First we'll rewrite into standard form: $f(x) = x^T \begin{bmatrix} 6 & 4 \\ 4 & 6 \end{bmatrix} x + x^T \begin{bmatrix} 5 \\ 6 \end{bmatrix} + 7$.

   We can take up to the smallest eigenvalue of F as the step size, as taking a step larger than that will lead to divergence rather than convergence. So we find the eigenvalues of F by solving $f(\lambda) = \lambda^2 - 12\lambda + 20$ to get $\lambda = 2, 10$.

   The max step size is then $\frac{2}{\lambda_{max}}$ (Thm 8.3). So any $0 < \alpha < \frac{2}{10}$ will be gaurenteed to converge.

10. **8.9 Consider the system of equations $h(x) = \begin{bmatrix} 4 + 3x_1 + 2x_2 \\ 1 + 2x_1 + 3x_2 \end{bmatrix}$**

    (a) Find when $h(x) = 0$.

    We can rewrite $h(x)$ as a matrix multiplication $h(x) = \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -4 \\ -1 \end{bmatrix}$ Then by multiplying by the inverse matrix we get $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} 3 & -2 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} -4 \\ -1 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$.

11. **8.10 Consider the function $f(x) = \frac{3}{2}(x_1^2 + x_2^2) + (1 + a)x_1x_2 - (x_1 + x_2) + b$**

    (a) $f(x) = \frac{1}{2}x^T \begin{bmatrix} 3 & 1+a \\ 1+a & 3 \end{bmatrix} x - x^T \begin{bmatrix} 1 \\ 1 \end{bmatrix} + b$

    (b) We can solve analytically for x^* as $x^* = Q^{-1}b = \frac{1}{9-(1+a)^2} \begin{bmatrix} 3 & -a-1 \\ -a-1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

    We need the determininet to be > 0 for positive definitite, so $-4 \leq a \leq 2$. The minimia always exists for any b.

    (c) If our alpha is $\frac{2}{5}$ then our max eigenvalue is 5. Solving for this yields $-3 \leq a \leq 1$.

12. **8.11 Consider $f(x) = \frac{1}{2}(x - c)^2$, minimized using an iterative algoritm**

    (a)

    $$f(x^{k+1}) = \frac{1}{2}(x^k - \alpha_k f'(x^k) - c)^2$$

    $$= \frac{1}{2}(x^k - c - \alpha_k f'(x^k))^2 = \frac{1}{2}[(x^k - c)^2 - 2\alpha_k f'(x^k)(x^k - c) + (\alpha_k f'(x^k))^2]$$

    $$= f(x^k) - \alpha_k f'(x^k)(x^k - c) + \frac{1}{2}(\alpha_k f'(x^k))^2$$

    Recall that $f'(x^k) = x^k - c$ so plugging in yields

    $$= f(x^k) - \frac{1}{2}\alpha_k^2(x^k - c)^2 = f(x^k)(1 - \alpha_k)^2 = f(x^{k+1})$$

    (b) Note that from a we get $f(x^k) = f(x^0)\prod_{i=0}^{k-1}(1 - \alpha_i)^2$

    So for this to converge $\prod_{i=0}^{\infty}(1 - \alpha_i)^2$ must go to 0. Since $0 < \alpha_i < 1$ this is the same as $\prod_{i=0}^{\infty}(1 - \alpha_i) = 0 \iff \sum_{k=0}^{\infty} \alpha_k = \infty$

13. **8.12 Consider minimizing $x^3 - x$**

    We can rewrite $f(x) = x(x+1)(x-1)$ and $\nabla f(x) = 3x^2 - 1$ so we know that the minima occurs at $\frac{1}{\sqrt{3}}$.

14. **8.15 Consider minimizing $\|ax - b\|^2$**

    (a) We can rewrite this as $f(x) = (ax-b)^T(ax-b) = (ax^T - b^T)ax - (ax^T - b^T)b = \|a\|^2x^2 - 2a^Tbx + \|b\|^2$.

    The minimia can be found when $\nabla f(x) = 0 \implies 2\|a\|^2x - 2a^Tb = 0 \implies x^* = \frac{a^Tb}{\|a\|^2}$

    (b) We can use Thm 8.3 to get that the max step size should be $\frac{2}{2\|a\|^2} = \frac{1}{\|a\|^2}$.

15. **8.16 Consider minimizing $\|Ax - b\|^2$**

    (a) We can rewrite this as

    $$f(x) = (Ax-b)^T(Ax-b) = (x^TA^T - b^T)Ax - (x^TA^T - b^T)b =$$

    $$= x^TA^TAx - b^TAx - x^TA^Tb + b^Tb = x^TA^TAx - 2b^TAx + b^Tb$$

    $\nabla f(x) = 2A^TAx - 2b^TA$ and $Q = 2A^TA$.

    (b)

    $$x^{k+1} = x^k - \alpha(\nabla f(x^k)) = x^k - \alpha 2(A^TAx - b^TA)$$

    (c) For $A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$ $A^TA = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$ and the eigenvalues of A^TA are 1, 4 respectively.

    So by Thm 8.3 this corresponds to a max step size of $\frac{1}{4}$.

16. **8.18 Show that steepest descent converges in one step iff $\nabla f(x^0)$ is an eigenvector.**

    For steepest descent to converge in one step

    $$x^1 = x^* = Q^{-1}b = x^0 + \alpha g^0 \implies b = Qx^0 + \alpha Qg^0 \implies Qx^0 - b = -\alpha Qg^0$$

    But $Qx^0 - b = g^0$ so $\frac{1}{\alpha}g^0 = Qg^0 \implies g^0$ is an eigenvector.
    Suppose g^0 was an eigenvector. Then we know $Qg^0 = \lambda g^0$ with $\lambda = \frac{1}{\alpha}$ (from above).

    $$x^1 = x^0 - \alpha g^0 \implies Qx^1 = Q(x^0 - \alpha g^0) \implies Qx^1 = Qx^0 - \alpha Qg^0 \implies Qx^1 = Qx^0 - g^0 = b$$

    but this means that $x^1 = Q^{-1}b = x^*$.

17. **8.21 Find the largest step size such that the algorithm is globally convergent**

    (a) We can find the eigenvalues of the hessian $Q = \begin{bmatrix} 6 & 4 \\ 4 & 6 \end{bmatrix}$. The max eigevalue is 10 so the max step size is $\frac{1}{5}$.

    (b) This is the same hessian as above $A + A^T = Q = \begin{bmatrix} 6 & 4 \\ 4 & 6 \end{bmatrix}$. The max eigevalue is 10 so the max step size is $\frac{1}{5}$.