Jesse Cai
MATH 164 Homework 2
304634445

1. **9.3 Consider minimizing** $f(x) = x^{\frac{4}{3}}$

   (a) We can find $\nabla f(x) = g(x) = \frac{4}{3}\sqrt[3]{x}$ and $f"(x) = \frac{4}{9}x^{-\frac{2}{3}}$

   $$x^{k+1} = x^k - F(x^k)^{-1}g^k = x^k - (\frac{4}{9}x^{-\frac{2}{3}})^{-1}\frac{4}{3}x^{\frac{1}{3}} = x^k - 3x^k = -2x^k$$

   (b) We can expand the fomula to get $x^k = -2^k x^0$ which does not converge to 0 except when $x^0 = 0$.

2. **9.4 Consider minimizing** $f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$

   (a) Note $f(x) \geq 0$, so the minima occurs when $f(x) = 0$. For this to be true, we can see that both $(1 - x_1)^2$ and $(100(x_2 - x_1^2)^2$ are 0. Solving the first equation yields $x_1 = 1$ and plugging this into the second part and solving yields $x_2 = 1$. So the unique minima is $[1, 1]$.

   (b) $\nabla f(x) = [400x_1^3 - 400x_1x_2 + 2x_1 - 2, 200(x_2 - x_1^2)]^T$

   $$F(x) = \begin{bmatrix} 1200x_1^2 - 400x_2 + 2 & -400x_1 \\ -400x_1 & 200 \end{bmatrix}$$

   $$F(x)^{-1} = \frac{1}{80000(x_1^2 - x_2) + 400} \begin{bmatrix} 200 & 400x_1 \\ 400x_1 & 1200x_1^2 - 400x_2 + 2 \end{bmatrix}$$

   $x_0 = [0, 0]^T$, $g_0 = [-2, 0]$, $x_1 = x_0 - F(x^0)^{-1}g^0 = -\frac{1}{400}\begin{bmatrix} 200 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} -2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

   $x_1 = [1, 0]^T$, $g_1 = [400, -200]$, $x_2 = x_1 - F(x^1)^{-1}g^1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} - \frac{1}{80400}\begin{bmatrix} 200 & 400 \\ 400 & 1202 \end{bmatrix}\begin{bmatrix} 400 \\ -200 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

   (c) $x_0 = [0, 0]^T$, $g_0 = [-2, 0]$, $x_1 = x_0 - 0.05g_0 = [0.1, 0]^T$
   $x_1 = [0.1, 0]^T$, $g_1 = [-1.4, 2]$, $x_2 = x_1 - 0.05g_1 = [0.17, 0.1]^T$

3. **10.1 Show that** $d^0 \dots d^{n-1}$ **are** Q **conjugate**

   We can do this via induction. Let $P(n) = \forall i < n : d_i^T Q d_n = 0$.

   Base Case: $d_0 = p_0$ and $d_1 = p_1 - \frac{p_1^T Q d_0}{d_0^T Q d_0} d_0$

   $$d_0^T Q d_1 = p_0^T Q(p_1 - \frac{p_1^T Q p_0}{p_0^T Q p_0} p_0) = p_0^T Q p_1 - \frac{p_1^T Q p_0}{p_0^T Q p_0} p_0^T Q p_0 = (p_1^T Q p_0)(1 - 1) = 0$$

   Inductive Step: Assume $P(n)$ holds, then we want to show $P(n) \implies P(n+1)$.

   $$d_{k+1} = p_{k+1} - \sum_{i=0}^{k} \frac{p_{k+1}^T Q d_i}{d_i^T Q d_i} d_i \implies d_j^T Q d_{k+1} = d_j^T Q p_{k+1} - \sum_{i=0}^{k} \frac{p_{k+1}^T Q d_i}{d_i^T Q d_i} d_i$$

   But we know by $P(n)$ that $i \neq j \implies d_j^T d_i = 0$ and then by symmetry of Q.

   $$d_j^T Q d_{k+1} = d_j^T Q p_{k+1} - \frac{p_{k+1}^T Q d_j}{d_j^T Q d_j} d_j^T Q d_j = d_j^T Q p_{k+1} - p_{k+1}^T Q d_j = 0$$

4. **10.2**

5. **10.3 Show that in the conjugate gradient method for a standard quadratic** $d_k^T Q d_k = -d_k^T Q g_k$

   Again we can do this via induction. Let $P(k) = d_k^T Q d_k = -d_k^T Q g_k$

   Base Case: By definition $g_0 = -d_0 \implies d_0^T Q d_0 = -d_0^T g_0$ so $P(0)$ holds.

   Inductive Step: Assume $d_k^T Q d_k = -d_k^T Q g_k$. By definition $d_{k+1} = -g_{k+1} - \beta_k d_k$

   Then since d_k, d_{k+1} are Q-conjugate:

   $$d_{k+1}^T Q d_{k+1} = d_{k+1}^T Q(-g_{k+1} - \beta_k d_k) = -d_{k+1}^T Q g_{k+1} - \beta_k d_{k+1}^T Q d_k = -d_{k+1}^T Q g_{k+1}$$

6. **10.4 Let Q be a real symmetric matrix**

   (a) Since Q is a real symmetric matrix, there exists an orthogonal basis of eigenvectors $\{v_1 \dots v_n\}$. This basis is $Q - conjugate$.
   Pick i, j s.t. $i \neq j$. Then because v_j is an eigenvector and since v_i, v_j are orthogonal.

   $$\implies v_i^T Q v_j = v^i \lambda_j v_j = \lambda_j v_i^T v_j = 0$$

   (b) If $\{d_1 \dots d_n\}$ is Q-conjugate and also orthogonal then.

   $$d_i^T Q d_j = 0 = d_i^T d_j = \lambda_j d_i^T d_j = d_i^T \lambda_j d_j \implies Q d_j = \lambda_j d_j$$

   So d_j must be an eigenvector.

7. **10.5 For a standard quadratic with $d_{k+1} = \gamma_k g_{k+1} + d_k$ find an expression for γ_k**
   Recall that our original definition of the conjugate gradient method give $d_{k+1} = -g_{k+1} + \beta_k d_k$.
   So

   $$\gamma_k g_{k+1} + d_k = -g_{k+1} + \beta_k d_k \implies \gamma_k = \frac{-1}{\beta_k} = \frac{d_k^T Q d_k}{g_{k+1}^T Q d_k}$$

8. **10.6**

9. **10.7 Consider a standard quadratic function $f(x)$ and $\phi(a) = f(x_0 + Da)$ where D is a matrix of rank r. Show ϕ is a standard quadratic**
   Note $f(x) = \frac{1}{2} x^T Q x - x^T b + c$

   $$\phi(a) = \frac{1}{2}(x_0 + Da)^T Q (x_0 + Da) - (x_0 + Da)^T b = x_0^T Q x_0 + (aD)^T Q D a - x_0^T b - (Da)^T b$$

   Using the fact that x_0 and b are constants, and then $(Da)^T = a^T D^T$ we get:

   $$\phi(a) = (Da)^T Q D a - (Da)^T b + c = a D^T D^t Q D a = a^t D b \implies Q' = D^T Q D, b' = Db$$

   Q' is positive definite.

10. **10.8 Consider a conjugate gradient algorithm applied to a quadratic function**

    (a) TODO

11. **10.9 Represent $f(x_1, x_2) = \frac{5}{2}x_1^2 + x_2^2 - 3x_1x_2 - x_2 - 7$ in standard form and find d_1.**

    $$f(x) = \frac{1}{2} x^T \begin{bmatrix} 5 & -3 \\ -3 & 2 \end{bmatrix} x - [0, 1]^T x - 7$$

12. **10.10 Consider minimizing $f(x_1, x_2) = \frac{5}{2}x_1^2 + \frac{1}{2}x_2^2 + 2x_1x_2 - 3x_1 - x_2$.**

    (a)

    $$f(x) = \frac{1}{2} x^T \begin{bmatrix} 5 & 2 \\ 2 & 1 \end{bmatrix} x - [3, 1]^T x$$

    (b) TODO

    (c) Solving this analyticaly, we get the same answeras above.

    $$x^* = Q^{-1} b = \begin{bmatrix} 1 & -2 \\ -2 & 5 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

13. **11.1**

14. **11.3**

15. **11.5**

16. **11.7**

17. **11.9**

18. **11.10**