Name: Uzair Akbar ID: 03697290

CONVEX OPTIMIZATION FOR ML & CV

Weekly Exercise 8
Exact Line Search

Exercise 1

The EVD of Q is given as

$$Q = U \Lambda U^T \Rightarrow Q^{-1} = U \Lambda^{-1} U^T$$

where

a) U is orthogonal, $\therefore U^{-1} = U^T$ [$\because Q$ symmetric]

b) $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ with eigenvalues $\lambda_i > 0$. [$\because Q$ positive definite]

Now, substituting the value of Q, we get

$$\frac{(x^T x)^2}{(x^T Q x)(x^T Q^{-1} x)} = \frac{\|x\|^4}{\langle x, U \Lambda U^T x \rangle \langle x, U \Lambda^{-1} U^T x \rangle}$$

$$= \frac{\|U^T x\|^4}{\langle U^T x, \Lambda U^T x \rangle \langle U^T x, \Lambda^{-1} U^T x \rangle}$$

[$\because \|x\| = \|Ux\|$ for orthogonal U.]

$$\Rightarrow \frac{(x^T x)^2}{(x^T Q x)(x^T Q^{-1} x)} = \frac{\|y\|^4}{\langle y, \Lambda y\rangle \langle y, \Lambda^{-1} y\rangle}$$

[substituting $y = U^T x$]

$$\Rightarrow \quad = \frac{\|y\|^4}{\left(\sum_{i=1}^{n} \lambda_i y_i^2\right)\left(\sum_{i=1}^{n} (1/\lambda_i) y_i^2\right)}$$

$$\Rightarrow \quad = \frac{1}{\left\{\sum_{i=1}^{n} \lambda_i \left(\frac{y_i}{\|y\|}\right)^2\right\}\left\{\sum_{i=1}^{n} \left(\frac{1}{\lambda_i}\right)\left(\frac{y_i}{\|y\|}\right)^2\right\}}$$

$$\Rightarrow \quad = \frac{1}{\left\{\sum_{i=1}^{n} \lambda_i z_i\right\}\left\{\sum_{i=1}^{n} (1/\lambda_i) z_i\right\}}$$

[substituting $z_i = \left(\frac{y_i}{\|y\|}\right)^2$
$\Rightarrow z_i \geq 0, \sum_i z_i = 1$]

$$\Rightarrow \quad = \frac{1}{\gamma \left\{\sum_{i=1}^{n} (1/\lambda_i) z_i\right\}}$$

[substituting $\gamma = \sum_{i=1}^{n} \lambda_i z_i$,
$\Rightarrow \gamma \in [\lambda_1, \lambda_n]$]

Now, we can get an upper bound on $\sum_{i=1}^{n} \left(1/\lambda_i\right) z_i$, as a simple interpolation of $[\lambda_1^{-1}, \lambda_n^{-1}]$ at γ:

$$\sum_{i=1}^{n} \left(1/\lambda_i\right) z_i \leq \frac{1}{\lambda_n} + \frac{1/\lambda_1 - 1/\lambda_n}{\lambda_n - \lambda_1}\left(\lambda_n - \gamma\right)$$

Therefore,

$$\Rightarrow \frac{(x^T x)^2}{(x^T Q x)(x^T Q^{-1} x)} \geq \frac{\gamma^{-1}}{\frac{1}{\lambda_n} + \frac{1/\lambda_1 - 1/\lambda_n}{\lambda_n - \lambda_1}(\lambda_n - \gamma)}$$

$$\geq \min_{\gamma \in [\lambda_n, \lambda_n]} \frac{\gamma^{-1}}{\frac{1}{\lambda_n} + \frac{1/\lambda_1 - 1/\lambda_n}{\lambda_n - \lambda_1}(\lambda_n - \gamma)}$$

simplification yields,

$$\Rightarrow \quad = \lambda_1 \lambda_2 \min_{\gamma} \frac{\gamma^{-1}}{(\lambda_1 + \lambda_n - \gamma)}$$

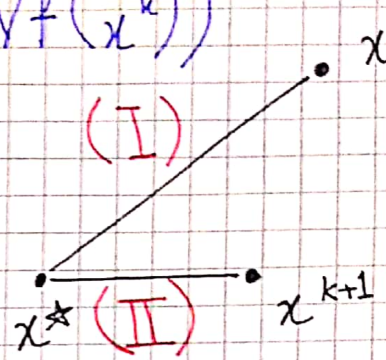
Solving for γ gives $\gamma^* = \frac{\lambda_1 + \lambda_n}{2}$. Substituting back gives the final result:

$$\therefore \frac{(x^T x)^2}{(x^T Q x)(x^T Q^{-1} x)} \geq \frac{4 \lambda_n \lambda_1}{(\lambda_n + \lambda_1)^2}$$

Exercise 2

For the line-search algorithm:

$$\because \tau^k = \operatorname{argmin}_\tau f(x^k - \tau \nabla f(x^k))$$



(i) —— $$\tau^k = \frac{\|\nabla f(x^k)\|^2}{\|\nabla f(x^k)\|^2_{\mathbb{Q}}}$$

Also:

$$\because \nabla f(x^k) = \mathbb{Q}x^k - b$$

(ii) —— $$\Rightarrow \quad = \mathbb{Q}(x^k - x^*)$$

[$\because$ for $x^k = x^*$
$\Rightarrow \nabla f(x^*) = \mathbb{Q}x^* - b = 0$
$\Rightarrow b = \mathbb{Q}x^*$]

We can now get

$$\because \|x^k - x^*\|^2_{\mathbb{Q}} = \langle x^k - x^*, \mathbb{Q}(x^k - x^*)\rangle$$

$$= \langle x^k - x^*, \mathbb{Q}\mathbb{Q}^{-1}\mathbb{Q}(x^k - x^*)\rangle$$

[multiplying ~~&~~ by $I = \mathbb{Q}\mathbb{Q}^{-1}$]

$$= \langle \mathbb{Q}^T(x^k - x^*), \mathbb{Q}^{-1}\mathbb{Q}(x^k - x^*)\rangle$$

$$= \langle \mathbb{Q}(x^k - x^*), \mathbb{Q}^{-1}\mathbb{Q}(x^k - x^*)\rangle$$

[$\because \mathbb{Q}$ symmitric]

$$= \|\mathbb{Q}(x^k - x^*)\|^2_{\mathbb{Q}^{-1}}$$

(I) $$\therefore \|x^k - x^*\|^2_{\mathbb{Q}} = \|\nabla f(x^k)\|^2_{\mathbb{Q}^{-1}}$$ —— (iii) [from (ii)]

Now,

$$\because \|x^k - x^\star\|_Q^2 - \|x^{k+1} - x^\star\|_Q^2 = \left\{ \begin{array}{l} \|x^k - x^\star\|_Q^2 \\ - \|x^k - \tau_k \nabla f(x^k) - x^\star\|_Q^2 \end{array} \right\}$$

$$\Rightarrow \quad = \left\{ \begin{array}{l} \left(\|x^k\|_Q^2 - 2\langle x^k, x^\star \rangle_Q + \|x^\star\|_Q^2 \right) \\ - \left(\begin{array}{l} \|x^k - \tau_k \nabla f(x^k)\|_Q^2 + \|x^\star\|_Q^2 \\ -2\langle x^k - \tau_k \nabla f(x^k), x^\star \rangle_Q \end{array} \right) \end{array} \right\}$$

$$\Rightarrow \quad = \|x^k\|_Q^2 - \|x^k - \tau_k \nabla f(x^k)\|_Q^2 - 2\tau_k \langle \nabla f(x^k), x^\star \rangle_Q$$

$$\Rightarrow \quad = \left\{ \begin{array}{l} 2\tau_k \langle x^k, \nabla f(x^k) \rangle_Q - 2\tau_k \langle \nabla f(x^k), x^\star \rangle_Q \\ - \tau_k^2 \|\nabla f(x^k)\|_Q^2 \end{array} \right\}$$

$$\Rightarrow \quad = 2\tau_k \langle \nabla f(x^k), x^k - x^\star \rangle_Q - \tau_k^2 \|\nabla f(x^k)\|_Q^2$$

$$\Rightarrow \quad = 2\frac{\|\nabla f(x^k)\|^4}{\|\nabla f(x^k)\|_Q^2} - \frac{\|\nabla f(x^k)\|^4}{\|\nabla f(x^k)\|_Q^2} \quad \left[\text{from (i)}\right]$$

$$\Rightarrow \quad = \frac{\|\nabla f(x^k)\|^4}{\|\nabla f(x^k)\|_Q^2} \quad \text{——— (iv)}$$

Finally,

$$\because \|x^{k+1}-x^*\|_Q^2 = \|x^k-x^*\|_Q^2 - \left(\|x^k-x^*\|_Q^2 - \|x^{k+1}-x^*\|_Q^2\right)$$

$$\Rightarrow \quad = \|x^k-x^*\|_Q^2 - \frac{\|\nabla f(x^k)\|^4}{\|\nabla f(x^k)\|_Q^2} \quad \text{[from (iv)]}$$

$$\Rightarrow \quad = \|\nabla f(x^k)\|_{Q^{-1}}^2 - \frac{\|\nabla f(x^k)\|^4}{\|\nabla f(x^k)\|_Q^2} \quad \text{[from (iii)]}$$

$$\Rightarrow \quad = \|\nabla f(x^k)\|_{Q^{-1}}^2 \left(1 - \frac{\|\nabla f(x^k)\|^4}{\|\nabla f(x^k)\|_Q^2 \cdot \|\nabla f(x^k)\|_{Q^{-1}}^2}\right)$$

$$\Rightarrow \quad = \|\nabla f(x^k)\|_{Q^{-1}}^2 \left(1 - \frac{(z^T z)^2}{(z^T Q z)(z^T Q^{-1} z)}\right)$$

[substituting $z = \nabla f(x^k)$]

$$\Rightarrow \quad \leq \|\nabla f(x^k)\|_{Q^{-1}}^2 \left(1 - \frac{4\lambda_1\lambda_n}{(\lambda_1+\lambda_n)^2}\right)$$

[exercise 1]

$$\Rightarrow \quad = \left(\frac{\lambda_1-\lambda_n}{\lambda_1+\lambda_n}\right)^2 \|\nabla f(x^k)\|_{Q^{-1}}^2$$

(II) $$\therefore \|x^{k+1}-x^*\|_Q^2 \leq \left(\frac{\lambda_1-\lambda_n}{\lambda_1+\lambda_n}\right)^2 \|x^k-x^*\|_Q^2 \quad \text{[from (iii)]}$$