Continuous Optimization: Assignment 3

Due on May 14, 2024

Honglu Ma **Hiroyasu Akada** **Mathivathana Ayyappan**

Exercise 1

Exercise 2

(a)

Since Q is a square matrix, we can write $Q = U\Lambda U^T$ where U is an orthogonal matrix and Λ is a diagonal matrix with the eigenvalues of Q i.e. λ_i on the diagonal. Then, we have

$$\begin{aligned}
\langle x, Qx\rangle &= \langle x, U\Lambda U^\top x\rangle \\
&= x^\top U\Lambda U^\top x \\
&= (U^\top x)^\top \Lambda U^\top x \\
&= (Ux)^\top \Lambda (Ux) && U = U^\top \\
&= \sum_{i=1}^{n} \lambda_i (Ux)_i^2 \\
&\leq \sum_{i=1}^{n} \lambda_{\max}(Q)(Ux)_i^2 && \lambda_i \leq \lambda_{\max}(Q) \\
&= \lambda_{\max}(Q) \sum_{i=1}^{n} (Ux)_i^2 \\
&= \lambda_{\max}(Q)(Ux)^\top Ux \\
&= \lambda_{\max}(Q)x^\top U^\top Ux \\
&= \lambda_{\max}(Q)x^\top x && U^\top U = I \\
&= \lambda_{\max}(Q)||x||^2
\end{aligned}$$

Similar derivation can be shown for the smallest eigenvalue: $\langle x, Qx\rangle \geq \lambda_{\min}(Q)||x||^2$.

(b)

Suppose λ is an eigenvalue of Q with eigenvector v. Then, we have

$$\begin{aligned}
Qv = \lambda v &\Rightarrow \tau Qv = \tau\lambda v \\
&\Rightarrow Iv - \tau Qv = Iv - \tau\lambda v \\
&\Rightarrow (I - \tau Q)v = (I - \tau \mathrm{diag}(\lambda))v
\end{aligned}$$

$I - \tau\mathrm{diag}(\lambda)$ is a matrix with same diagonal entries $1 - \tau\lambda$

$$\Rightarrow (I - \tau Q)v = (1 - \tau\lambda)v$$

Above shows that $1 - \tau\lambda$ is an eigenvalue of $I - \tau Q$.

$$\begin{aligned}
&\Rightarrow (I - \tau Q)(I - \tau Q)v = (I - \tau Q)(1 - \tau\lambda)v \\
&\Rightarrow (I - \tau Q)(I - \tau Q)v = (1 - \tau\lambda)(I - \tau Q)v \\
&\Rightarrow (I - \tau Q)(I - \tau Q)v = (1 - \tau\lambda)(1 - \tau\lambda)v \\
&\Rightarrow (I - \tau Q)^2 v = (1 - \tau\lambda)^2 v
\end{aligned}$$

Exercise 2 continued on next page. . .

Thus $(1-\tau\lambda)^2$ is an eigenvalue of $(I-\tau Q)^2$ for each eigenvalue λ of Q.

Exercise 3

(a)

Projection of $\mathbf{v}$ onto the space spanned by the columns of $\mathbf{A}$ means that we want to find a point in the column space of $\mathbf{A}$ that is closest to $\mathbf{v}$ i.e. for vector $\mathbf{p}$ in the column space we want the distance between $\mathbf{v}$ and $\mathbf{p}$ to be as small as possible

$$\operatorname*{argmin}_p \operatorname{dist}(\mathbf{v}, \mathbf{p})$$

which is equivalent to

$$\operatorname*{argmin}_p ||\mathbf{v} - \mathbf{p}||$$

or

$$\operatorname*{argmin}_p \sqrt{\langle \mathbf{v} - \mathbf{p}, \mathbf{v} - \mathbf{p} \rangle}$$

(b)

In case of $m = 1$, the subspace spanned by $\mathbf{u}$ is just a line $c\mathbf{u}$ for some constant $c \in \mathbb{R}$. We have

$$\operatorname*{argmin}_c \sqrt{\langle \mathbf{v} - c\mathbf{u}, \mathbf{v} - c\mathbf{u} \rangle}$$

which by definition of the inner product, is equivalent to

$$\operatorname*{argmin}_c (\mathbf{v} - c\mathbf{u})^\top \mathbf{Q} (\mathbf{v} - c\mathbf{u})$$

To find minimum, we take derivative of function

$$f(c) = (\mathbf{v} - c\mathbf{u})^\top \mathbf{Q} (\mathbf{v} - c\mathbf{u}) = \mathbf{v}^\top \mathbf{Q} \mathbf{v} - 2(\mathbf{v}^\top \mathbf{Q} \mathbf{u})c + (\mathbf{u}^\top \mathbf{Q} \mathbf{u})c^2$$

We have

$$f'(c) = 2(\mathbf{u}^\top \mathbf{Q} \mathbf{u})c - 2(\mathbf{v}^\top \mathbf{Q} \mathbf{u})$$

and the second derivative

$$f''(c) = 2(\mathbf{u}^\top \mathbf{Q} \mathbf{u}) \geq 0$$

which indicate function f is convex now we just have to solve $f'(c) = 0$

$$2(\mathbf{u}^\top \mathbf{Q} \mathbf{u})c - 2(\mathbf{v}^\top \mathbf{Q} \mathbf{u}) = 0$$
$$c = \frac{\mathbf{v}^\top \mathbf{Q} \mathbf{u}}{\mathbf{u}^\top \mathbf{Q} \mathbf{u}}$$

and the projection $\mathbf{p} = \frac{\mathbf{v}^\top \mathbf{Q} \mathbf{u}}{\mathbf{u}^\top \mathbf{Q} \mathbf{u}} \mathbf{u}$

(c)

In case of $m > 1$, we have a vector $\hat{\mathbf{c}} \in \mathbb{R}^m$ defined by

$$\hat{\mathbf{c}} := \begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix}$$

where the projection $\mathbf{p} = \mathbf{A}\hat{\mathbf{c}}$. We have

$$\underset{\hat{\mathbf{c}}}{\operatorname{argmin}} \, (\mathbf{v} - \mathbf{A}\hat{\mathbf{c}})^\top \mathbf{Q}(\mathbf{v} - \mathbf{A}\hat{\mathbf{c}})$$

To find the minimum, similarly, we define a function

$$g(\hat{\mathbf{c}}) = \mathbf{v}^\top \mathbf{Q}\mathbf{v} - 2\mathbf{v}^\top \mathbf{Q}\mathbf{A}\hat{\mathbf{c}} + \hat{\mathbf{c}}^\top \mathbf{A}^\top \mathbf{Q}\mathbf{A}\hat{\mathbf{c}}$$

We calculate its gradient

$$\nabla g(\hat{\mathbf{c}}) = -2\mathbf{v}^\top \mathbf{Q}\mathbf{A} + 2\mathbf{A}^\top \mathbf{Q}\mathbf{A}\hat{\mathbf{c}}$$

and its Hessian

$$\nabla^2 g(\hat{\mathbf{c}}) = 2\mathbf{A}^\top \mathbf{Q}\mathbf{A} = \mathbf{H}$$

We can show that $\mathbf{H}$ is positive semi-definite. Suppose λ is an eigenvalue of $\mathbf{H}$ and with corresponding eigenvector $\mathbf{e}$

$$\begin{aligned} \mathbf{H}\mathbf{e} &= \lambda \mathbf{e} \\ 2\mathbf{A}^\top \mathbf{Q}\mathbf{A}\mathbf{e} &= \lambda \mathbf{e} \\ 2\mathbf{e}^\top \mathbf{A}^\top \mathbf{Q}\mathbf{A}\mathbf{e} &= \lambda \mathbf{e}^\top \mathbf{e} \\ \lambda = \frac{2(\mathbf{A}\mathbf{e})^\top \mathbf{Q}\mathbf{A}\mathbf{e}}{\mathbf{e}^\top \mathbf{e}} &\geq 0 \end{aligned}$$

Thus we can get the minimizer $\hat{\mathbf{c}}$ by solving $\nabla g(\hat{\mathbf{c}}) = 0$

$$\begin{aligned} -2\mathbf{v}^\top \mathbf{Q}\mathbf{A} + 2\mathbf{A}^\top \mathbf{Q}\mathbf{A}\hat{\mathbf{c}} &= 0 \\ \mathbf{A}^\top \mathbf{Q}\mathbf{A}\hat{\mathbf{c}} &= \mathbf{v}^\top \mathbf{Q}\mathbf{A} \end{aligned}$$