Exercise 3

$\boldsymbol{A}$ is a matrix of size $m \times n$. Define $\boldsymbol{P} = \boldsymbol{A}^T\boldsymbol{A}$ (size: $n \times n$) and $\boldsymbol{Q} = \boldsymbol{A}\boldsymbol{A}^T$ (size: $m \times m$).

Part (a)

For any vector $\boldsymbol{y}$ of size $n \times 1$:

$$\boldsymbol{y}^T\boldsymbol{P}\boldsymbol{y} = \boldsymbol{y}^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{y} = (\boldsymbol{A}\boldsymbol{y})^T\boldsymbol{A}\boldsymbol{y} = ||\boldsymbol{A}\boldsymbol{y}||^2 \geq 0$$

Hence proved.
For any vector $\boldsymbol{z}$ of size $m \times 1$:

$$\boldsymbol{z}^T\boldsymbol{Q}\boldsymbol{z} = \boldsymbol{z}^T\boldsymbol{A}\boldsymbol{A}^T\boldsymbol{z} = (\boldsymbol{A}^T\boldsymbol{z})^T\boldsymbol{A}^T\boldsymbol{z} = ||\boldsymbol{A}^T\boldsymbol{z}||^2 \geq 0$$

Hence proved.

Let λ be an eigenvalue of $\boldsymbol{P}$ with non-zero eigenvector $\boldsymbol{v}$. Then:

$$\boldsymbol{P}\boldsymbol{v} = \lambda\boldsymbol{v} \longrightarrow \boldsymbol{A}\boldsymbol{A}^T\boldsymbol{v} = \lambda\boldsymbol{v}$$

Pre-multiplying both sides by $\boldsymbol{v}^T$:

$$\boldsymbol{v}^T\boldsymbol{A}\boldsymbol{A}^T\boldsymbol{v} = \lambda\boldsymbol{v}^T\boldsymbol{v} \longrightarrow (\boldsymbol{A}^T\boldsymbol{v})^T\boldsymbol{A}^T\boldsymbol{v} = \lambda\boldsymbol{v}^T\boldsymbol{v} \longrightarrow ||\boldsymbol{A}^T\boldsymbol{v}||^2 = \lambda||\boldsymbol{v}||^2 \longrightarrow \lambda = \frac{||\boldsymbol{A}^T\boldsymbol{v}||^2}{||\boldsymbol{v}||^2} \geq 0$$

Similarly, let η be an eigenvalue of $\boldsymbol{Q}$ with non-zero eigenvector $\boldsymbol{w}$. Then:

$$\boldsymbol{Q}\boldsymbol{w} = \eta\boldsymbol{w} \longrightarrow \boldsymbol{A}^T\boldsymbol{A}\boldsymbol{w} = \eta\boldsymbol{w}$$

Pre-multiplying both sides by $\boldsymbol{w}^T$:

$$\boldsymbol{w}^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{w} = \eta\boldsymbol{w}^T\boldsymbol{w} \longrightarrow (\boldsymbol{A}\boldsymbol{w})^T\boldsymbol{A}\boldsymbol{w} = \eta\boldsymbol{w}^T\boldsymbol{w} \longrightarrow ||\boldsymbol{A}\boldsymbol{w}||^2 = \eta||\boldsymbol{w}||^2 \longrightarrow \eta = \frac{||\boldsymbol{A}\boldsymbol{w}||^2}{||\boldsymbol{w}||^2} \geq 0$$

Hence, proved.

Part (b)

Let $\boldsymbol{u}$ be an eigenvector of $\boldsymbol{P}$ with eigenvalue λ. Then:

$$\boldsymbol{P}\boldsymbol{u} = \lambda\boldsymbol{u} \longrightarrow \boldsymbol{A}^T\boldsymbol{A}\boldsymbol{u} = \lambda\boldsymbol{u}$$

Pre-multiplying both sides by $\boldsymbol{A}$, we have:

$$\boldsymbol{A}\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{u} = \lambda\boldsymbol{A}\boldsymbol{u} \longrightarrow (\boldsymbol{A}\boldsymbol{A}^T)\boldsymbol{A}\boldsymbol{u} = \lambda\boldsymbol{A}\boldsymbol{u} \longrightarrow \boldsymbol{Q}\boldsymbol{A}\boldsymbol{u} = \lambda\boldsymbol{A}\boldsymbol{u}$$

Let $\boldsymbol{v} = \boldsymbol{A}\boldsymbol{u}$, then:

$$\boldsymbol{Q}\boldsymbol{v} = \lambda\boldsymbol{v}$$

Therefore, $\boldsymbol{A}\boldsymbol{u}$ is an eigenvector of $\boldsymbol{Q}$ with eigenvalue λ.

Similarly, let $\boldsymbol{v}$ be an eigenvector of $\boldsymbol{Q}$ with eigenvalue μ. Then:

$$\boldsymbol{Q}\boldsymbol{v} = \mu\boldsymbol{v} \longrightarrow \boldsymbol{A}\boldsymbol{A}^T\boldsymbol{v} = \mu\boldsymbol{v}$$

Pre-multiplying both sides by $\boldsymbol{A}^T$, we have:

$$\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{A}^T\boldsymbol{v} = \mu\boldsymbol{A}^T\boldsymbol{v} \longrightarrow (\boldsymbol{A}^T\boldsymbol{A})\boldsymbol{A}^T\boldsymbol{v} = \mu\boldsymbol{A}^T\boldsymbol{v} \longrightarrow \boldsymbol{P}\boldsymbol{A}^T\boldsymbol{v} = \mu\boldsymbol{A}^T\boldsymbol{v}$$

Let $\boldsymbol{u} = \boldsymbol{A}^T\boldsymbol{v}$, then:

$$\boldsymbol{P}\boldsymbol{u} = \mu\boldsymbol{u}$$

Therefore, $\boldsymbol{A}^T\boldsymbol{v}$ is an eigenvector of $\boldsymbol{P}$ with eigenvalue μ.

Size of vector $\boldsymbol{u}$ would be $n \times 1$ and size of vector $\boldsymbol{v}$ would be $m \times 1$.

Part (c)

Since $\boldsymbol{v}_i$ is an eigenvector of $\boldsymbol{Q}$ (with an eigenvalue $\lambda_i \geq 0$, as proved in the part (a)), we have:

$$\boldsymbol{Q}\boldsymbol{v}_i = \lambda_i \boldsymbol{v}_i \longrightarrow \boldsymbol{A}\boldsymbol{A}^T\boldsymbol{v}_i = \lambda_i \boldsymbol{v}_i$$

Dividing both sides by $||\boldsymbol{A}^T\boldsymbol{v}_i||_2$, we have:

$$\frac{\boldsymbol{A}\boldsymbol{A}^T\boldsymbol{v}_i}{||\boldsymbol{A}^T\boldsymbol{v}_i||_2} = \frac{\lambda_i \boldsymbol{v}_i}{||\boldsymbol{A}^T\boldsymbol{v}_i||_2}$$

As given in the question, $\boldsymbol{u}_i = \frac{\boldsymbol{A}^T\boldsymbol{v}_i}{||\boldsymbol{A}^T\boldsymbol{v}_i||_2}$. Substituting this in the original equation:

$$\frac{\boldsymbol{A}\boldsymbol{A}^T\boldsymbol{v}_i}{||\boldsymbol{A}^T\boldsymbol{v}_i||_2} = \frac{\lambda_i \boldsymbol{v}_i}{||\boldsymbol{A}^T\boldsymbol{v}_i||_2} \longrightarrow \boldsymbol{A}\boldsymbol{u}_i = \frac{\lambda_i}{||\boldsymbol{A}^T\boldsymbol{v}_i||_2}\boldsymbol{v}_i$$

Therefore, $\gamma_i = \frac{\lambda_i}{||\boldsymbol{A}^T\boldsymbol{v}_i||_2}$.

The eigenvalue λ_i has to be non-negative as proved in part (a) and the denominator is definitely positive (it cannot be zero since $\boldsymbol{u}_i$ won't be well defined then), so γ_i is also non-negative.

Part (d)

Define $\boldsymbol{U} = [\boldsymbol{v}_1|\boldsymbol{v}_2|\boldsymbol{v}_3|....|\boldsymbol{v}_m]$, $\boldsymbol{V} = [\boldsymbol{u}_1|\boldsymbol{u}_2|\boldsymbol{u}_3|....|\boldsymbol{u}_m]$ and a $m \times n$ diagonal matrix $\boldsymbol{\Gamma}$ with diagonal entries γ_1, γ_2 ...., γ_m. From the previous part, we know $\boldsymbol{A}\boldsymbol{u}_i = \gamma_i\boldsymbol{v}_i$. Thus we have:

$$\boldsymbol{A}\boldsymbol{V} = [\boldsymbol{A}\boldsymbol{u}_1|\boldsymbol{A}\boldsymbol{u}_2|\boldsymbol{A}\boldsymbol{u}_3|....|\boldsymbol{A}\boldsymbol{u}_m] = [\gamma_1\boldsymbol{v}_1|\gamma_2\boldsymbol{v}_2|\gamma_3\boldsymbol{v}_3|....|\gamma_m\boldsymbol{v}_m] = \boldsymbol{U}\boldsymbol{\Gamma}$$

Post-multiplying both sides by $\boldsymbol{V}^T$, we have:

$$\boldsymbol{A}\boldsymbol{V}\boldsymbol{V}^T = \boldsymbol{U}\boldsymbol{\Gamma}\boldsymbol{V}^T \longrightarrow \boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Gamma}\boldsymbol{V}^T$$

where the last equality follows from the fact that $\boldsymbol{V}$ is an orthonormal matrix (since $\boldsymbol{u}_i^T\boldsymbol{u}_j = 0$ for $i \neq j$), thus implying that $\boldsymbol{V}\boldsymbol{V}^T = \boldsymbol{V}^T\boldsymbol{V} = \boldsymbol{I}$.
Hence, proved.