Let $A \in \mathbb{R}^{m \times n}$ and $B \equiv A^T A \in \mathbb{R}^{n \times n}$.

$B$ is symmetric and positive semidefinite &
(i.e. $x^T B x \geq 0 \;\; \forall x \in \mathbb{R}^n$)
which imply that all the eigenvalues of $B$ are greater or equal to $0$. Moreover, by the Spectral Theorem, $B$ is diagonalizable by means of a unitary matrix $Q$ (i.e. $Q^T = Q^{-1}$)

$\lambda_i \geq 0$ because

$$\boxed{B = Q^T \Lambda Q} \quad \text{where } \Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$$

Since $Q$ is unitary we also have $\boxed{\|Qx\|_2 = \|x\|_2 \;\; \forall x \in \mathbb{R}^n}$

**Proposition ①**

Let $B \equiv A^T A \in \mathbb{R}^{n \times n}$, then $\forall x, y \in \mathbb{R}^n$ we have

$$x^T B y \leq \lambda_{max}(B) \, \|x\|_2 \, \|y\|_2$$

$\lambda_{max}(B) \equiv \max\{\lambda_1, \ldots, \lambda_n\}$

**Proof:**

$$x^T B y = x^T Q^T \Lambda Q y = (Qx)^T \Lambda Q y$$

we denote $(Qx)_i$ and $(Qy)_i$ the $i^{th}$ element of these vectors then we have



$$x^T B y = (Qx)^T \Lambda Q y = \sum_{i=1}^{n} \lambda_i (Qx)_i (Qy)_i$$

$$\leq \lambda_{max}(B) \sum_{i=1}^{n} (Qx)_i (Qy)_i$$

$$\leq \lambda_{max}(B) \sqrt{\sum_{i=1}^{n} (Qx)_i^2} \sqrt{\sum_{i=1}^{n} (Qy)_i^2}$$

Cauchy-Schwarzz inequality

$$= \lambda_{max}(B) \, \|x\|_2 \, \|y\|_2$$

\#

**Corollary:**

Let $B \equiv A^T A \in \mathbb{R}^{n \times n}$ then $\forall x \in \mathbb{R}^n$ we have

$$x^T B x \leq \lambda_{max}(B) \, \|x\|_2^2$$

**Proposition ②**

$$\|B\|_2 = \lambda_{max}(B) \quad \text{for } B \equiv A^T A \in \mathbb{R}^{n \times n}$$

(true for any positive semidefinite any symmetric matrix)



**Proof:**

not sure if it's the definition but it's true!

By definition (?)

$$\|B\|_2 = \sqrt{\lambda_{max}(B^T B)}$$

But $B^T B = B^2 = Q^T \Lambda \underbrace{Q Q^T}_{Id} \Lambda Q = Q^T \Lambda^2 Q$

which implies that $\lambda_{max}(B^2) = \lambda_{max}(B)^2$
(since $\lambda_i \geq 0$)

Therefore, we have

$$\|B\|_2 = \sqrt{\lambda_{max}(B^T B)} = \sqrt{\lambda_{max}(B)^2} = \lambda_{max}(B)$$

\#

Combing these 3 results, we have that for all $A \in \mathbb{R}^{m \times n}$ and all $x \in \mathbb{R}^n$

$$\|Ax\|_2^2 = x^T B x \leq \lambda_{max}(B) \, \|x\|_2^2 = \|B\|_2 \, \|x\|_2^2$$

Hence $\boxed{\|Ax\|_2^2 \leq \|B\|_2 \, \|x\|_2^2}$

Moreover $\|B\|_F = \sqrt{Tr(B^T B)} = \sqrt{Tr(B^2)} \geq \sqrt{\lambda_{max}(B)^2}$
$= \lambda_{max}(B) = \|B\|_2$

Thus, $\boxed{\|Ax\|_2^2 \leq \|B\|_F \, \|x\|_2^2}$ where $\|B\|_F \equiv \sqrt{\sum_{i,j} B_{i,j}^2}$