Let $V$ and $W$ be normed vector spaces and let $T : V \to W$ be a linear transformation. The operator norm of $T$ is

$$
\|T\| = \sup \left\{ \frac{|T(v)|_W}{|v|_V} : v \in V, v \neq 0 \right\}
$$

It is easy to check that the operator norm is in fact a norm. The operator norm is a measure of the maximum amount that $T$ stretches vectors in $V$. In the case that $V$ and $W$ are Euclidean spaces, the following theorem provides a simple way to calculate the operator norm.

**Theorem 1.** Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be linear with matrix $A$. Then $\|T\| = \sqrt{\lambda_{\max}}$, where $\lambda_{\max}$ is the largest eigenvalue of the matrix $A^tA$.

**Example 1.** Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the transformation with matrix

$$
A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}.
$$

Then

$$
A^tA = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 10 & 14 \\ 14 & 20 \end{bmatrix}
$$

The eigenvalues of $A^tA$ are the roots of $\lambda^2 - 30\lambda + 4$, so the largest eigenvalue is

$$
\lambda_{\max} = \frac{30 + \sqrt{884}}{2} = 15 + \sqrt{221}
$$

so $\|T\| = \sqrt{15 + \sqrt{221}}$.

The proof of Theorem 1 relies on the Spectral Theorem.

**Theorem 2.** (Spectral Theorem) Let $B$ be an $n \times n$ symmetric matrix. Then there exists an orthonormal basis $\{v_1, v_2, \ldots, v_n\}$ for $\mathbb{R}^n$ consisting of eigenvectors of $B$. That is

$$
\langle v_i, v_j \rangle = \begin{cases} 
1 & i = j \\
0 & i \neq j
\end{cases}
$$

and $Bv_j = \lambda_j v_j$.

**Proof of Theorem 1.** Let $B = A^tA$. Then $B$ is a symmetric matrix since $B^t = (A^tA)^t = A^t(A^t)^t = A^tA = B$. For any vector $v \in \mathbb{R}^n$,

$$
|Av|^2 = \langle Av, Av \rangle = \langle A^tAv, v \rangle = \langle Bv, v \rangle.
$$
By the Spectral Theorem, $B$ has an orthonormal eigenbasis $\{v_1, v_2, \ldots, v_n\}$, so $v$ can be written as a linear combination

$$v = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n$$

of the eigenvectors of $B$. Thus

$$\langle Bv, v \rangle = \langle B(c_1 v_1 + c_2 v_2 + \cdots + c_n v_n), c_1 v_1 + c_2 v_2 + \cdots + c_n v_n \rangle$$

$$= \langle \lambda_1 c_1 v_1 + \lambda_2 c_2 v_2 + \cdots + \lambda_n c_n v_n, c_1 v_1 + c_2 v_2 + \cdots + c_n v_n \rangle$$

$$= \lambda_1 c_1^2 + \lambda_2 c_2^2 \cdots + \lambda_n c_n^2$$

and

$$|v|^2 = \langle v, v \rangle = \langle c_1 v_1 + c_2 v_2 + \cdots + c_n v_n, c_1 v_1 + c_2 v_2 + \cdots + c_n v_n \rangle$$

$$= c_1^2 + c_2^2 + \cdots + c_n^2.$$