2-Norm

Proposition 2.7.5. If $A \in \mathbf{R}^{m\times n}$, then

$$\|A\|_2 = \sqrt{\max_{\mu\in\lambda(A^TA)} \mu},$$

i.e., $\| A\|_2$ is the square root of the largest eigenvalue of A^TA .

Proof. To calculate $\| A\|_2$, we find first $\| A\|_2^{\,2}$. Thus,

$$\|A\|_2 = \max_{\|\mathbf{x}\|_2=1} \|A\mathbf{x}\|_2 \Rightarrow \ \| A\|_2^{\,2} = \max_{\|\mathbf{x}\|_2=1} \|A\mathbf{x}\|_2^{\,2} = \max_{\mathbf{x}^T\mathbf{x}=1} \mathbf{x}^T A^T A\mathbf{x}.$$

Let $A^TA = B \in \mathbf{R}^{n\times n}$. The matrix B is a symmetric matrix because

$$B^T = (A^TA)^T = A^TA$$

and

$$\mathbf{x}^TA^TA\mathbf{x} = \mathbf{x}^TB\mathbf{x} = \begin{bmatrix} \xi_1 & \cdots & \xi_n \end{bmatrix} \begin{bmatrix} b_{1\,;\,1} & \cdots & b_{1\,;\,n} \\ \vdots & \ldots & \vdots \\ b_{n\,;\,1} & \cdots & b_{n\,;\,n} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix} =$$

$$= \begin{bmatrix} \xi_1 & \cdots & \xi_n \end{bmatrix} \begin{bmatrix} b_{1\,;\,1}\xi_1 + \ldots + b_{1\,;\,n}\xi_n \\ \cdots \quad \cdots \quad \cdots \\ b_{n\,;\,1}\xi_1 + \ldots + b_{n\,;\,n}\xi_n \end{bmatrix} =$$

$$= \left[\xi_1 \sum_{j=1}^n b_{1;\,j}\xi_j + \ldots + \xi_n \sum_{j=1}^n b_{n;\,j}\xi_j\right],$$

$$\mathbf{x}^T\mathbf{x} = \sum_{j=1}^n \xi_j^{\,2},$$

then $\mathbf{x}^TA^TA\mathbf{x}$ is a function of n variables $\xi_1, \ldots, \xi_n$ and

$$\frac{\partial\left(\mathbf{x}^TA^TA\mathbf{x}\right)}{\partial\xi_i} = \sum_{j=1}^n b_{i;\,j}\xi_j + \sum_{j=1}^n b_{j;\,i}\xi_j \overset{b_{i;\,j}=b_{j;\,i}}{=} 2\sum_{j=1}^n b_{i;\,j}\xi_j = 2\left[A^TA\mathbf{x}\right]_i,$$

$$\frac{\partial(\,\mathbf{x}^T\mathbf{x})}{\partial\xi_i} = 2\xi_i = 2\left[\mathbf{x}\right]_i.$$

The problem of finding $\max_{\mathbf{x}^T\mathbf{x}=1} \mathbf{x}^TA^TA\mathbf{x}$ is a problem of finding the relative extremum. To solve our problem we form an auxiliary function

$$\Phi(\xi_1;\, \ldots, \xi_n; \rho) = \mathbf{x}^TA^TA\mathbf{x} + \mu\left(1 - \mathbf{x}^T\mathbf{x}\right).$$

To find the stationary points of Φ, we form the system of equations:

$$\frac{\partial \Phi}{\partial \xi_i} = 0 \ \ (\text{i} = 1 : \text{n}) \ \ \wedge \ \ \frac{\partial \Phi}{\partial \rho} = 0,$$

i.e.,

$$\begin{cases} 2\left[A^T A\mathbf{x}\right]_i - 2\mu\, [\mathbf{x}]_i = 0 \ \ (\text{i} = 1 : \text{n}) \\ 1 - \mathbf{x}^T\mathbf{x} = \mathbf{0} \end{cases}$$

or

$$\begin{cases} A^T A\mathbf{x} = \mu\mathbf{x}\,, \\ \|\mathbf{x}\|_2 = 1\,. \end{cases}$$

Thus, any stationary point for relative extremum is the normed vector $\mathbf{x}$ corresponding to an eigenvalue of $A^T A$. Let us express from the relation $A^T A\mathbf{x} = \mu\mathbf{x}$ the eigenvalue μ. We obtain that $\mu = \mathbf{x}^T A^T A\mathbf{x}$, where $\|\mathbf{x}\|_2 = 1$. Comparing this result with the original formula for finding $\|A\|_2^2$, we notice that $\|A\|_2^2 = \max\limits_{\mu \in \lambda(A^T A)} \mu$. Thus,

$$\|A\|_2 = \sqrt{\max_{\mu \in \lambda(A^T A)} \mu},$$

i.e., $\|A\|_2$ is the square root of the largest eigenvalue of $A^T A$.

Corollary 2.7.1. If matrix $A \in \mathbf{R}^{m \times n}$ is symmetric, then

$$\|A\|_2 = \max_{\lambda \in \lambda(A)} \lambda.$$