Norms of Matrices

We can measure matrix sizes using vector norms, because $\mathbb{R}^{m\times n}$ is a vector space. Although the names are different, all of the p-norms above give matrix norms if the matrix is stretched into one long vector. The only matrix norm of this flavor that is used often is the Frobenius norm

$$\|A\|_F \;=\; (\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2)^{1/2} \;=\; \mathrm{tr}(A^tA)^{1/2}.$$

It is the vector 2-norm applied to a stretched out version of the matrix. Any norm on the vector space $\mathbb{R}^{mn}$ can be used as a matrix norm on $\mathbb{R}^{m\times n}$.

But matrices are also operators, and as such we can measure their size by how much they stretch the vectors on which they operate:

$$\|A\|_{\mathcal{D},\mathcal{R}} = \max_{\|x\|_{\mathcal{D}}=1}\|Ax\|_{\mathcal{R}}.$$

You can think of it as the radius of the smallest ball centered at the origin of $\mathbb{R}^m$ that contains the image of the ball of radius 1 centered at the origin of $\mathbb{R}^n$ (the *unit ball* in $\mathbb{R}^n$). This is the norm *induced* by (or subordinate to) the vector norms $\|\cdot\|_D$ and $\|\cdot\|_R$. For some authors, this is the *only* type of matrix norm.

With some algebra, we can see that

$$\|A\|_1 = \max_{1\le j\le n}\sum_{i=1}^{m}|a_{ij}|, \qquad \text{(the max |column sum|), and}$$

$$\|A\|_\infty = \max_{1\le i\le m}\sum_{j=1}^{m}|a_{ij}|, \qquad \text{(the max |row sum|).}$$

Another twist that matrices add to the theory of norms is function composition (matrix multiplication). Is the norm of the product related to the product of the norms? Three norms (or two or even one) are called *consistent*, or *submultiplicative*, if they satisfy

$$\|AB\|_{lr} \le \|A\|_l\|B\|_r.$$

This always holds with the induced norms. That is, if $\|A\|$ is induced from $\|.\|_D$ and $\|.\|_R$, then by definition (the max part, combined with $\|\alpha x\| = |\alpha|\|x\|$)

$$\|Ax\|_R \le \|A\|\|x\|_D.$$

The induced norm is the smallest matrix norm satisfying this submultiplicitivity. Transformations from one inner-product space to another have a natural norm, being the norm induced by the vector norms induced by the respective inner products. The matrix 2-norm is of this flavor, and it has many nice properties, but it is difficult to compute. One way to approximate it is to notice that it is the square root of the largest eigenvalue of A^tA (maximize x^tA^tAx over unit circle):

$$\|A\|_2 = \sqrt{\lambda_1(A^tA)}.$$