21. Vector and Matrix Norms

In order to measure the size of a number, disregarding its sign, we use absolute value. In order to measure the size of a vector, ignoring direction, there are various norms we can use, including but not limited to the standard length.

A vector norm is a function $\| \cdot \| : \mathbb{R}^k \rightarrow \mathbb{R}$ satisfying

(a) $\| x \| > 0$ if $x \neq 0$, and $\| 0 \| = 0$,
(b) $\| kx \| = |k| \| x \|$ for $k \in \mathbb{R}$ (i.e., $k$ a scalar),
(c) $\| x + y \| \leq \| x \| + \| y \|$.

The standard norms which we will use are the “$\ell_p$” norms for $p = 1, 2$ and $\infty$:

$\| x \|_1 = |x_1| + \cdots + |x_n|$  
$\| x \|_2 = \sqrt{x_1^2 + \cdots + x_n^2}$  
$\| x \|_\infty = \max_{1 \leq i \leq n} |x_i|$.

A matrix norm satisfies

(a’) $\| A \| > 0$ if $A \neq 0$, and $\| 0 \| = 0$,
(b’) $\| kA \| = |k| \| A \|$ for $k \in \mathbb{R}$,
(c’) $\| A + B \| \leq \| A \| + \| B \|$,
(d’) $\| AB \| \leq \| A \| \| B \|$.

Standard examples are:

$\| A \|_1 = \max_{1 \leq j \leq n} \left( \sum_{i=1}^{n} |a_{ij}| \right) = \text{maximum absolute column sum}$

$\| A \|_2 = \sqrt{\lambda}$ where $\lambda$ is the largest eigenvalue of $A^tA$

$\| A \|_E = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}^2}$

$\| A \|_\infty = \max_{1 \leq i \leq n} \left( \sum_{j=1}^{n} |a_{ij}| \right) = \text{maximum absolute row sum}$

These norms are pairwise compatible, i.e.,

$\| Ax \|_1 \leq \| A \|_1 \| x \|_1$

$\| Ax \|_2 \leq \| A \|_2 \| x \|_2$

$\| Ax \|_2 \leq \| A \|_E \| x \|_2$

$\| Ax \|_\infty \leq \| A \|_\infty \| x \|_\infty$
22. Types of Error

Let $x_c$ be the vector of calculated values of $x$, while $x_t$ denotes the true solution to $Ax = b$. Then, using one of the above vector norms,

\[
\begin{align*}
||x_c - x_t|| & \quad \text{is the absolute error;} \\
\frac{||x_c - x_t||}{||x_t||} & \quad \text{is the relative error;} \\
Ax_c - b & \quad \text{is the residual (a vector);} \\
\frac{||Ax_c - b||}{||b||} & \quad \text{is the relative residual (a number).}
\end{align*}
\]

Of course, different norms give different values for the errors.

23. Residual Correction

One type of error is fairly easy to detect and correct. If the relative residual is significantly large, then:

(a) Calculate $r = Ax_c - b$ in double precision.
(b) Solve $Ae = r$.
(c) Let $x = x_c - e$.

Then

\[
Ax = A(x_c - e) = Ax_c - Ae = Ax_c - r = b.
\]

The catch in the above calculation is that it assumes that $Ae = r$ was solved precisely, while in fact it also probably contains some error. However, $r$ is usually rather small compared to $b$, making the absolute error in the second calculation small compared to the original, and hence $x$ should be closer to the true solution than $x_c$. If $e$ is very large, it’s a good idea to repeat this process.

24. Ill-Conditioned Matrices

If $A$ has a small eigenvalue (so that in some sense it is almost singular), then the error can be large even though the residual is small. For if $Av = \lambda v$ with $\lambda << 1$, then $A(x_t + v) = b + \lambda v$, so that the right-hand side is close to $b$ even though $v$ may be relatively large. This is a more severe problem because it is hard to detect.

25. Error Estimates

A primitive estimate of the error in solving $Ax = b$ is given by

\[
||x_c - x_t|| \leq ||A^{-1}|| \cdot ||r||
\]

where $r = Ax_c - b$ is the residual vector. We may obtain a lower bound for $||A^{-1}||$ by noting that for any vector $v$, we have

\[
\frac{||v||}{||Av||} \leq ||A^{-1}||.
\]
The _standard estimate_ is given by

$$\frac{\|x_c - x_t\|}{\|x_t\|} \leq \|A\| \|A^{-1}\| \|r\| \|b\|.$$