Improving on Gaussian Elimination

Gaussian elimination is the standard technique used to solve systems of linear equations. The big problem with Gaussian elimination is its inefficiency. For very large n, it is impractical to apply an $O(n^3)$ algorithm to solve an n by n system.

In chapter 7 we are going to study various alternatives to Gaussian elimination. Most of these schemes are iterative schemes that start with a starting guess and apply an iterative process to construct a sequence of vectors that we hope converge toward a solution.

The first technical issue we need to deal with in chapter 7 is the question of convergence. How do we formally express the fact that a sequence of vectors is converging to a limit?

Measuring distances

Before we can discuss convergence, we need to start with some reasonable notion of distance. That is, given two vectors $\mathbf{x}$ and $\mathbf{y}$, how far apart are $\mathbf{x}$ and $\mathbf{y}$? The following definition formalizes the notion of distance applied to vectors.

Definition A *vector norm* on $\mathbb{R}^n$ is a function, $\| \|$, from $\mathbb{R}^n$ to $\mathbb{R}$ with the following properties:

1. $\|\mathbf{x}\| \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$.
2. $\|\mathbf{x}\| = 0$ if and only if $\mathbf{x} = 0$.
3. $\|\alpha\, \mathbf{x}\| = |\alpha|\, \|\mathbf{x}\|$ for all $\alpha \in \mathbb{R}$ and $\mathbf{x} \in \mathbb{R}^n$.
4. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ for all $\mathbf{x},\mathbf{y} \in \mathbb{R}^n$.

In practice, several different norms are used to measure distances in $\mathbb{R}^n$. The l_∞ norm is defined by

$$\|\mathbf{x}\|_\infty = \max_{1 \leq j \leq n} |x_j|$$

The l_1 norm is defined by

$$\|\mathbf{x}\|_1 = \sum_{j=1}^{n} |x_j|$$

The l_2 norm is defined by

$$\|\mathbf{x}\|_2 = \sqrt{\sum_{j=1}^{n} |x_j|^2}$$

The l_2 norm conforms most closely to the standard notion of distance in $\mathbb{R}^n$ and is the norm most commonly used.

With a notion of distance in place, we can move on to a definition of convergence.

Definition A sequence $\{\mathbf{x}^{(k)}\}$ of vectors in $\mathbb{R}^n$ is said to converge to $\mathbf{x}$ with respect to the l_2 norm if

$$\lim_{k\to\infty} \|\mathbf{x}^{(k)} - \mathbf{x}\|_2 = 0$$

It is possible to show that if a sequence converges with respect to any of the three norms shown above then it converges with respect to all the other norms.

Matrix Norms

A closely related concept is the idea of a matrix norm.

Definition A *matrix norm* on the set of all n by n matrices is a real valued function $\| \ \|$ on the set of n by n matrices satisfying for all n by n matrices A and B and all real numbers α:

1. $\|A\| \geq 0$
2. $\|A\| = 0$ if and only if A is a matrix with all 0 entries.
3. $\|\alpha A\| = |\alpha| \, \|A\|$
4. $\|A + B\| \leq \|A\| + \|B\|$
5. $\|A B\| \leq \|A\| \, \|B\|$

Matrix norms are most commonly constructed via the following theorem.

Theorem If $\| \ \|$ is a vector norm on $\mathbb{R}^n$, then

$$\|A\| = \max_{\|\mathbf{x}\|=1} \|A x\|$$

is a matrix norm.