The Residual

In section 7.3 we learned that we can solve some linear systems

$$A\,x = b$$

by systematically transforming them into fixed point problems

$$x = T\,x + c$$

and then generating an interated sequence of vectors $\mathbf{x}^{(k)}$ by doing

$$\mathbf{x}^{(k)} = T\,\mathbf{x}^{(k-1)} + \mathbf{c}$$

As we generate this sequence of vectors we would like to know how close we are getting to a limit. A simple way to measure our progress is to note that when we reach a limiting value $\mathbf{x}$ we should have a vector that satisfies the original system:

$$A\,x = b$$

A simple way to measure the "goodness" of an approximation $\mathbf{x}^{(k)}$ is to compute $A\,\mathbf{x}^{(k)}$ and see how far from $\mathbf{b}$ that is. That idea motivates the following definition.

Definition The *residual* associated with a candidate solution vector $\mathbf{x}^{(k)}$ is

$$\mathbf{r}^{(k)} = \mathbf{b} - A\,\mathbf{x}^{(k)}$$

Relaxation Methods

Now that we have an objective measure of how good or bad an approximate solution $\mathbf{x}^{(k)}$ is, we can develop methods that target this quantity and seek to shrink it as quickly as possible. For reasons that will be clear in a little while, these methods are known as *relaxation methods*.

Consider the Gauss-Seidel method.

$$(D - L)\,\mathbf{x}^{(k)} = U\,\mathbf{x}^{(k-1)} + \mathbf{b}$$

Consider a particular row in the $\mathbf{x}^{(k)}$ vector. Here is how we would go about computing that entry:

$$D_{i,i}\left(\mathbf{x}^{(k)}\right)_i = \sum_{j=1}^{i-1} L_{i,j}\left(\mathbf{x}^{(k)}\right)_j + \sum_{j=i+1}^{n} U_{i,j}\left(\mathbf{x}^{(k-1)}\right)_j + \mathbf{b}_i$$

It turns out that lurking somewhere in this equation is something that can be made to resemble a residual. The trick is to start by adding and subtracting a factor of $D_{i,i}\left(\mathbf{x}^{(k-1)}\right)_i$ on the right hand side.

$$D_{i,i}\left(\mathbf{x}^{(k)}\right)_i = D_{i,i}\left(\mathbf{x}^{(k-1)}\right)_i + \left(\sum_{j=1}^{i-1} L_{i,j}\left(\mathbf{x}^{(k)}\right)_j\right) - D_{i,i}\left(\mathbf{x}^{(k-1)}\right)_i + \left(\sum_{j=i+1}^{n} U_{i,j}\left(\mathbf{x}^{(k-1)}\right)_j\right) + \mathbf{b}_i$$

Recalling that $A = D - L - U$, the terms after the first term on the right can be rearranged to look like a residual:

$$D_{i,i} \left(\mathbf{x}^{(k)}\right)_i = D_{i,i} \left(\mathbf{x}^{(k-1)}\right)_i + \left(\mathbf{b}_i - A\, \tilde{\mathbf{x}}_i\right)$$

Where $\tilde{\mathbf{x}}_i$ is the i^{th} *intermediate vector between* $\mathbf{x}^{(k-1)}$ and $\mathbf{x}^{(k)}$

$$\tilde{\mathbf{x}}_i = (\left(\mathbf{x}^{(k)}\right)_1, \left(\mathbf{x}^{(k)}\right)_2, \left(\mathbf{x}^{(k)}\right)_3, \ldots, \left(\mathbf{x}^{(k)}\right)_{i-1}, \left(\mathbf{x}^{(k-1)}\right)_i, \ldots, \left(\mathbf{x}^{(k-1)}\right)_n)$$

Consider now the form

$$\left(\mathbf{x}^{(k)}\right)_i = \left(\mathbf{x}^{(k-1)}\right)_i + \frac{\mathbf{b}_i - A\, \tilde{\mathbf{x}}_i}{D_{i,i}}$$

This expression takes the form

$$\left(\mathbf{x}^{(k)}\right)_i = \left(\mathbf{x}^{(k-1)}\right)_i + (\text{error term})$$

The formula suggests that we get from $\mathbf{x}^{(k-1)}$ to $\mathbf{x}^{(k)}$ by computing and adding an error term. This process of moving toward a result by computing and adding error terms goes by the name of *relaxation*. What if we could accelerate this relaxation process by cranking up the error term, in effect *overrelaxing*?

$$\left(\mathbf{x}^{(k)}\right)_i = \left(\mathbf{x}^{(k-1)}\right)_i + \omega\, (\text{error term})$$

Under some circumstances, this can actually work. To work out the details we have to work backwards from this form to something that resembles the original formula for the Gauss-Seidel method.

$$D_{i,i} \left(\mathbf{x}^{(k)}\right)_i = D_{i,i} \left(\mathbf{x}^{(k-1)}\right)_i + \omega \left(\left(\sum_{j=1}^{i-1} L_{i,j} \left(\mathbf{x}^{(k)}\right)_j \right) - D_{i,i} \left(\mathbf{x}^{(k-1)}\right)_i + \left(\sum_{j=i+1}^{n} U_{i,j} \left(\mathbf{x}^{(k-1)}\right)_j \right) + \mathbf{b}_i \right)$$

This is row i of a more general formula

$$D\, \mathbf{x}^{(k)} = (1-\omega)\, D\, \mathbf{x}^{(k-1)} + \omega\, L\, \mathbf{x}^{(k)} + \omega\, U\, \mathbf{x}^{(k-1)} + \omega\, \mathbf{b}$$

or

$$\mathbf{x}^{(k)} = \left(D - \omega\, L\right)^{-1} \left((1-\omega)\, D + \omega\, U\right) \mathbf{x}^{(k-1)} + \omega \left(D - \omega\, L\right)^{-1} \mathbf{b} = T_\omega\, \mathbf{x}^{(k-1)} + \mathbf{c}_\omega$$

When $\omega > 1$ this is known as the method of *successive overrelaxation*, or the SOR method for short.