Iterative Methods II - Choosing splittings

We now see that we can define an iterative method based on any splitting

$$A = M - K, \quad R = M^{-1}K, \quad c = M^{-1}b$$

and it will converge if $\rho(R) < 1$. This gives us two goals:

1) M^{-1} should be easy to compute
2) $\rho(R)$ should be small.

We establish notation. If A has no zeros on the diagonal, let

$$A = D - \tilde{L} - \tilde{U} = D(I - L - U)$$

where $-\tilde{L}$ and $-\tilde{U}$ are the strictly upper and lower triangle lar parts of A, $D L = \tilde{L}$ and $D U = \tilde{U}$.

Jacobi's Method.

Choose $A = D - (\tilde{L} + \tilde{U})$. Then $R = D^{-1}(\tilde{L} + \tilde{U})$, $c = D^{-1}b$. {This corresponds to solving for x_j in the jth equation using whatever the current values are for other variables.} To see this

$$x_{n+1,j} = \left(D^{-1}(\tilde{L} + \tilde{U})\, x_n + D^{-1}b\right)_j$$

$$= \frac{1}{d_{jj}}\left(\sum_{k \neq j} (\tilde{L} + \tilde{U})_{jk}\, x_{n,k} + b_j\right)$$

$$= \frac{1}{a_{jj}}\left(-\sum_{k \neq j} a_{jk}\, x_{n,k} + b_j\right).$$

Now you can assume that having updated $(j-1)$ elements of x_{n+1} already, it would improve things to use these to compute $x_{n+1,j}$ rather than the $x_{n,k}$ for $k \in 1, .., j-1$.

This leads to the Gauss-Seidel method. We write it first term-by-term

$$x_{m+1,j} = \frac{1}{a_{jj}}\left(b_j - \underbrace{\sum_{k=1}^{j-1} a_{jk} x_{m+1,k}}_{\text{new guys}} - \underbrace{\sum_{k=j+1}^{n} a_{jk} x_{m,k}}_{\text{older guys.}} \right)$$

We would like to write this in our standard form. We start with

$$a_{jj}\, x_{m+1,j} = b_j - \sum_{k=1}^{j-1} a_{jk} x_{m+1,k} - \sum_{k=j+1}^{n} a_{jk} x_{m,k}$$

or

$$\sum_{k=1}^{j} a_{jk} x_{m+1,j} = -\sum_{k=j+1}^{n} a_{jk} x_{m,k} + b_j .$$

or

$$(D - \tilde{L})\, x_{m+1} = \tilde{U} x_m + b$$

so

$$x_{m+1} = (D-\tilde{L})^{-1} \tilde{U} x_m + (D-\tilde{L})^{-1} b.$$
$$= \left(D(I-L)\right)^{-1} (DU)^{-1} x_m + (D(I-L))^{-1} b$$

$$= (I-L)^{-1} D^{-1} \tilde{U} \cancel{D^{-1}}\, x_m + (I-L)^{-1} D^{-1} b$$

$$= (I-L)^{-1} \mathcal{U}\, x_m + (I-L)^{-1} D^{-1}\, b.$$

This gives the Gauss-Seidel splitting:

$$A = \underbrace{D(I-L)}_{M} - \underbrace{\tilde{\mathcal{U}}}_{K}$$

The next method we introduce is called successive overrelaxation, and denoted SOR(ω) where ω is a parameter.

We will use (instead of all "new" ~~of~~ GS) a linear combination of ~~all "old" Jacobi a linear~~ new and old values.

As before, this looks more motivated when we write it as a sum:

$$X_{m+1,j} = \underbrace{(1-\omega)\,x_{m,j}}_{\text{old value}} + \underbrace{\frac{\omega}{a_{jj}}\left(b_j - \sum_{k=1}^{j-1} a_{jk}\,x_{m+1,k} - \sum_{k=j+1}^{n} a_{jk}\,x_{m,k}\right)}_{\text{Gauss-Seidel value}}$$

Rearranging, we have

$$a_{jj}\,x_{m+1,j} + \omega\sum_{k=1}^{j-1} a_{jk}\,x_{m+1,k} = (1-\omega)\,a_{jj}\,x_{m,j} - \omega\sum_{k=j+1}^{n} a_{jk}\,x_{m,k} + \omega b_j$$

we then have

$$(D - \omega\tilde{L})\,x_{m+1} = \left((1-\omega)D + \omega\tilde{U}\right)x_m + \omega b$$

or

$$x_{m+1} = \left(D(I-\omega L)\right)^{-1}\left((1-\omega)D + \omega\tilde{U}\right)x_m + \omega\left(D(I-\omega L)\right)^{-1} b.$$

$$= (I-\omega L)^{-1}D^{-1}\left((1-\omega)D + \omega\tilde{U}\right)x_m + \omega(I-\omega L)^{-1}D^{-1}b.$$

$$= (I - \omega L)^{-1} \left((1-\omega) I + \omega U \right) x_m$$

$$+ \omega (I - \omega L)^{-1} D^{-1} b.$$

This is the splitting

$$\cancel{A = (D - \omega \tilde{L})^{-1} - (1-\omega) D}$$

$$\cancel{A = (D - \omega \tilde{L}) - ((1-\omega) \tilde{L}}$$

$$M = (D - \omega \tilde{L})$$

$$M^{-1} K = (D - \omega \tilde{L})^{-1} \left((1-\omega) D + \omega \tilde{U} \right)$$

$$\omega A = (D - \omega \tilde{L}) - \left((1-\omega) D + \omega \tilde{U} \right)$$

$$= -\omega D - \omega \tilde{L} - \omega \tilde{U}$$

or

$$A = \underbrace{\frac{1}{\omega} (D - \omega \tilde{L})}_{M} - \underbrace{\frac{1}{\omega} \left((1-\omega) D + \omega \tilde{U} \right)}_{K}.$$

How should we pick ω? Well, if Gauss-Seidel is pushing the solution in a good direction, it might help to go further ($\omega > 1$). This is called over relaxation.

Next time: convergence criteria and speed for Jacobi, GS, and SOR(ω).