Conjugate Gradients

Positive Definite $A = A'$. To solve $Az = b$ for $z = A^{-1}b$, choose $x_{-1} := x_0$ arbitrarily and, for $n = 0, 1, 2, 3, \dots$ in turn, compute $x_{n+1} := x_n + ß_n\cdot(b - Ax_n) + \mu_n\cdot(x_n - x_{n-1})$ with scalars $ß_n$ and μ_n chosen to minimize $(x_{n+1}-z)'A(x_{n+1}-z) = x_{n+1}'\cdot Ax_{n+1} - 2b'x_{n+1} + z'Az$. Simplify typography by dropping subscripts from $r_n := b - Ax_n$, $d_n := x_n - x_{n-1}$, $ß_n$ and μ_n to find that

$$[r, d]'A[r, d]\begin{bmatrix}\beta\\ \mu\end{bmatrix} = \begin{bmatrix}r'Ar & r'Ad\\ d'Ar & d'Ad\end{bmatrix}\begin{bmatrix}\beta\\ \mu\end{bmatrix} = \begin{bmatrix}r'r\\ d'r\end{bmatrix} = [r, d]'r .$$

Solve this equation for $ß$ and μ . At least one solution always exists if A is positive definite.

In the absence of roundoff, successive residuals $r_n := b - Ax_n$ turn out to be orthogonal; in fact

$$r_m'r_n = 0 = d_m'Ad_n \quad \text{for all } m > n \ge 0 .$$

Consequently $r_m = o$ and $x_m = z$ at least as soon as m equals the dimension of A . But the point of the iteration is *not* to iterate that many times when the dimension is huge. Instead, take advantage of the tendency of residuals r_n and increments d_n to dwindle as n increases, and stop iterating when they both become small enough.

Overrelaxation

Positive Definite $A = A' = -L + V - L'$ in which $V := \mathrm{Diag}(A) = V'$ is Positive Definite too, and $-L = \mathrm{Subdiag}(A)$. Given A and b we seek $z := A^{-1}b$. Starting from an arbitrary initial guess x_0 , for $n = 0, 1, 2, 3, \dots$ in turn, ordinary Gauss-Seidel iteration solves $V(x_{n+1}-x_n) = b + Lx_{n+1} - Vx_n + L'x_n$ for $x_{n+1} = (V - L)^{-1}(b + L'x_n)$. Then $(x_{n+1}-z) = E(x_n-z)$ where $E = (V - L)^{-1}L'$ can be shown to have eigenvalues all with magnitudes less than 1 , though not necessarily much less unless $\|L\|$ is rather smaller than the smallest eigenvalue of V .

To accelerate convergence, consider using an Over/Underrelaxation parameter δ confined to $-1 < \delta < 1$ for reasons to be explained later. To solve $Az = b$ for $z = A^{-1}b$, choose x_0 arbitrarily and, for $n = 0, 1, 2, 3, \dots$ in turn, solve $V(x_{n+1}-x_n) = (1+\delta)(b + Lx_{n+1} - Vx_n + L'x_n)$ for $x_{n+1} = (V - (1+\delta)L)^{-1}\big(Vx_n + (1+\delta)(b - Vx_n + L'x_n)\big)$. Then $(x_{n+1}-z) = E(x_n-z)$ where $E = (V - (1+\delta)L)^{-1}\big(-\delta V + (1+\delta)L'\big)$. Note that, because L lies strictly below the diagonal, (product of all E's eigenvalues) $= \det(E) = \det(-\delta I)$. Therefore at least one eigenvalue of E has magnitude at least as big as $|\delta|$. This is why we keep $-1 < \delta < 1$.

E and its eigenvalues depend upon δ , as well as A , but the dependence is obscure except in special cases. An important special case arises when $A = \begin{bmatrix} I & -B' \\ -B & I\end{bmatrix}$. In this case every eigenvalue α of A has the form $\alpha = 1 \pm ß$ where $ß$ is either a singular value of B or, if B is not square, 0 . And in this case every eigenvalue ε of E can be shown easily to satisfy $(\varepsilon+\delta) = \pm ß(1+\delta)\sqrt{\varepsilon}$;

$$\varepsilon = \Big((1+\delta)ß/2 \pm \sqrt{\big(((1+\delta)ß/2)^2 - \delta \big)} \Big)^2 .$$

The largest of the magnitudes of eigenvalues ε is minimized when $\delta = \big(\|B\|/\big(1 + \sqrt{(1-\|B\|^2)}\big)\big)^2$, and then every eigenvalue ε has the same magnitude δ .