Lesson 7

CSPP58001 Numerical Methods:
TA: Kyle Gerard Felker

February 21, 2013

1 Detour: Krylov subspace methods

While you will not be responsible for knowing this information for the homework , it is worth introducing since these methods are among the fastest in the modern day. Define

$$\kappa_j = \{\mathbf{b}, A\mathbf{b}, \ldots, A^{j-1}\mathbf{b}\}$$

the Krylov subspace. Note that these products are easy to compute. Let's form the matrix with the corresponding column vectors

$$K_j = \begin{bmatrix}\mathbf{b} & A\mathbf{b} & \ldots & A^{j-1}\mathbf{b}\end{bmatrix}$$

We show the connection to iterative methods by returning to the general formula

$$x^{k+1} = (I - A)x^k + b$$

Starting with a guess $x^{(0)} = 0$, we get

$$x^{(1)} = b$$

$$x^{(2)} = 2b - Ab$$

$$x^{(3)} = 3b - Ab + A^2b$$

Question: Can we find a better basis in κ_j that gets us to our solution much faster?

What we mean by "better" selects the particular Krylov method:

Conjugate gradient (CG): We make the residual better

$$r_j \equiv \mathbf{b} - A\mathbf{x}_j$$

is made orthogonal to the basis κ_i for all $i < j$. This only works robustly for symmetric positive definite matrices (SDP). These are symmetric matrices $A = A^T$ such that

$$\mathbf{x}^T A\mathbf{x} > 0$$

for all vectors.

Step 1: Start with an ugly, ill-conditioned Krylov subspace and make them orthonormal.

$$Q = [\mathbf{q}_1, \ldots, \mathbf{q}_n]$$

where $\mathbf{q}_i^T \mathbf{q}_j = 0$ for $i \neq j$
$\mathbf{q}_i^T \mathbf{q}_j = 1$ for $i = j$
You can do the Gram-Schmidt method, but this is highly inefficient. Here is our algorithm (Arnoldi iteration). First, let

$$\mathbf{q}_1 = \frac{\mathbf{b}}{||\mathbf{b}||}$$

---

for $j = 1, \ldots, n-1$ **do**
    $t = Aq_j$ ▷ next Krylov entry
    for $i = 1, \ldots, j$ **do** ▷ project onto all previous basis entries
        $h_{ij} = q_i^T t$
        $t = t - h_{ij} q_i$
    end for
    $h_{i+1,j} = ||t||$
    $q_{j+1} = \frac{t}{h_{i+1,j}}$
end for

With the above definitions for the entries of a matrix H, and using the $\mathbf{q}_i$ as the columns of matrix Q, we get an equivalence

$$AQ = \begin{bmatrix} A\mathbf{q}_1 & \ldots & A\mathbf{q}_n \end{bmatrix} = QH$$

This factorization allows us to use the orthonormality of Q ($Q^T Q = I$) to conclude

$$Q^T AQ = H$$

If A is symmetric, then $Q^T AQ$ is symmetric as well $\Rightarrow H^T = H$. But, we know that H is upper Hessenberg (nearly triangular), so H is thus tridiagonal.

1.1 example

Consider

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}$$

with $\mathbf{b} = (1, 1, 1, 1)$. Forming the Krylov matrix,

$$\kappa_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \\ 1 & 3 & 9 & 27 \\ 1 & 4 & 16 & 64 \end{bmatrix}$$

A check of the Krylov matrix's condition number in MATLAB shows that this is ill-conditioned, so this is a bad basis for our problem. From Arnoldi iteration algorithm, we begin with $\mathbf{q}_1 = (0.5, 0.5, 0.5, 0.5)$ $A\mathbf{q}_1 = (0.5, 1, 1.5, 2)$