

CAAM 551 Problem set 5

Due Monday 12 Nov.

Problem 1

Derivation of Bi-CGStab

Consider the problem solve $Ax = b$ for x with $A \in \mathbb{R}^{n \times n}$ and $b \in \mathbb{R}^n$. There are two relevant spaces

$$1) \mathcal{V} = \mathcal{K}_k(A, b) \quad \text{and} \quad 2) \mathcal{W} = \mathcal{K}_k(A^T, c)$$

where $c \in \mathbb{R}^n$ is arbitrary but not orthogonal to b ($c^T b \neq 0$, usually we take $c = b$).

FOM: Find $x \in \mathcal{V}$ such that

$$w^T(b - Ax) = 0, \quad \text{for all } w \in \mathcal{W}$$

Suppose $V, W \in \mathbb{R}^{n \times k}$ are both rank k with $\text{Range}(V) = \mathcal{V}$ and $\text{Range}(W) = \mathcal{W}$.

Note: If $\hat{A} \equiv W^T A V$ is nonsingular then the unique solution to FOM is given by $x = V y$ with $\hat{A} y = \hat{b}$ and $\hat{b} = W^T b$.

- Qx i) Prove: If \hat{A} is singular then no other choice of bases \tilde{V}, \tilde{W} will produce a nonsingular $\tilde{W}^T A \tilde{V}$.
- Qx ii) Let H_k be the $k \times k$ Hankel matrix whose i, j -th entry is $H_k(i, j) = c^T A^{i+j-1} b$ for $i, j = 1, 2, \dots, k$. Prove that FOM has a unique solution x_k if and only if H_k is non singular.

If P_j, Q_j are built sequentially so that $\text{Range}(P_j) = \mathcal{K}_j(A, b)$ and $\text{Range}(Q_j) = \mathcal{K}_j(A^T, c)$ for $j = 1, 2, \dots, k$ and we put $P_j = [p_0, p_1, \dots, p_{j-1}]$ and $Q_j = [q_0, q_1, \dots, q_{j-1}]$ then $p_i = \phi_i(A)b$ and $q_i = \psi_i(A^T)c$ for $i = 0, 1, 2, \dots, j-1$ where ϕ_i, ψ_i are polynomials of degree i . We intend to derive a flexible class of iterative methods via polynomial recursion: We shall generate polynomial sequences ψ_j, ϕ_j, π_j defined by the recursions

$$\begin{aligned}\psi_0 &= \phi_0 = \pi_0 = 1 \\ \psi_{j+1}(t) &= \psi_j(t) - \omega_j t \psi_j(t), \\ \phi_{j+1}(t) &= \phi_j(t) - \alpha_j t \pi_j(t), \\ \pi_{j+1}(t) &= \phi_{j+1}(t) + \beta_j \pi_j(t),\end{aligned}$$

for $j = 0, 1, 2, \dots$

Then, we define

$$q_j = \psi_j(A^T)c, \quad r_j = \phi_j(A)b, \quad p_j = \pi_j(A)b,$$

and from the polynomial recursions, we derive the vector recursions

$$q_{j+1} = q_j - A^T q_j \omega_j, \quad r_{j+1} = r_j - A p_j \alpha_j, \quad p_{j+1} = r_{j+1} + p_j \beta_j \quad \text{for } j = 0, 1, 2, \dots$$

with $q_0 = c$, $r_0 = b$, $p_0 = r_0$.

The coefficients α_j and β_j are constructed to propagate the Bi-Conjugate Direction conditions

$$q_i^T r_j = 0, \quad q_i^T A p_j = 0 \quad \text{for all } 0 \leq i < j.$$

The formulas for α_j, β_j are

$$\alpha_j = \frac{q_j^T r_j}{q_j^T A p_j} \quad \text{and} \quad \beta_j = \frac{q_{j+1}^T r_{j+1} \alpha_j}{q_j^T r_j \omega_j}.$$

- Qx iii) Prove that this choice of α_j, β_j propagates the two bi-conjugacy conditions by showing $q_i^T r_j = 0$ and $q_i^T A p_j = 0$ for $0 \leq i < j$ implies $q_i^T r_{j+1} = 0$ and $q_i^T A p_{j+1} = 0$ for $0 \leq i < j + 1$.
(Hint: use $\alpha_j q_j^T A p_j = q_j^T r_j$ and $\omega_j A^T q_j = q_j - q_{j+1}$.)
- Qx iv) Let $K_j \equiv [b, Ab, \dots, A^{j-1}b]$ and let $\tilde{K}_j \equiv [c, A^T c, \dots, (A^T)^{j-1}c]$ both have rank j . Show $P_j = K_j R_j$ and $Q_j = \tilde{K}_j \tilde{R}_j$ where both R_j and \tilde{R}_j are nonsingular and upper triangular. Note: You needn't give the elements of R_j and \tilde{R}_j explicitly. Just explain why such a relationship must exist and why the two R-matrices must be upper triangular.
- Qx v) Prove $Q_j^T A P_j = D_j$ is lower triangular and that $\tilde{R}_j^T H_j R_j = D_j$. Thus $H_j = (L_j D_j) U_j$ with $L_j^T = \tilde{R}_j^{-1}$ and $U_j = R_j^{-1}$, providing an LU-Factorization of H_j *without pivoting for stability!* What are the diagonal elements of D_j ? (Note: H_j is defined in Qx ii)

Qx vi) Let the sequence x_j be defined through the recursion $x_0 = 0$, $x_{j+1} = x_j + p_j \alpha_j$ for $j = 1, 2, \dots$. Prove that in exact arithmetic, this sequence is identical to the sequence of iterates produced by the Bi-Conjugate gradient algorithm (Saad, Algorithm 7.3. BCG, p. 223) *regardless of the choice of the ω_j* . Prove that this sequence is well defined up to step k if and only if H_j is nonsingular for $1 \leq j \leq k$. What is the relation between the singularity of H_j and the breakdown of the Bi-Lanczos and BCG algorithms?

Up to this point all we have accomplished is to construct an alternative version of the BCG algorithm. However, we have yet to determine the ω_j . We now want to put these to good use. We are going to avoid the need for A^T and also to make a useful choice of the ω_j . Instead of propagating $r_j = \phi_j(A)b$ and $p_j = \pi_j(A)b$, we shall propagate $\hat{r}_j = \psi_j(A)\phi_j(A)b$ and $\hat{p}_j = \psi_j(A)\pi_j(A)b$.

Qx vii) Derive the vector recurrences for these new sequences. Show

$$\begin{aligned}\hat{r}_{j+1} &= (I - \omega_j A)(\hat{r}_j - A\hat{p}_j\alpha_j) \\ \hat{p}_{j+1} &= \hat{r}_{j+1} + (I - \omega_j A)\hat{p}_j\beta_j\end{aligned}$$

Thus the sequences may be propagated by

$$\begin{aligned}s_j &= \hat{r}_j - A\hat{p}_j\alpha_j \\ \hat{r}_{j+1} &= s_j - As_j\omega_j \\ \hat{p}_{j+1} &= \hat{r}_{j+1} + (\hat{p}_j - A\hat{p}_j\omega_j)\beta_j\end{aligned}$$

We shall choose ω_j to minimize $\|\hat{r}_{j+1}\|$.

Qx viii) Recalling that $q_0 = c$, $r_0 = b$, $p_0 = r_0$, verify the following formulas:

$$\alpha_j = \frac{q_0^T \hat{r}_j}{q_0^T A \hat{p}_j}, \quad \omega_j = \frac{s_j^T A s_j}{\|A s_j\|^2}, \quad \beta_j = \frac{q_0^T \hat{r}_{j+1}}{q_0^T \hat{r}_j} \frac{\alpha_j}{\omega_j}.$$

Finally, we must give a recursion for the sequence \hat{x}_j which we will define via $\hat{r}_j = b - A\hat{x}_j$.

Qx ix) Use $\hat{r}_{j+1} = b - A\hat{x}_{j+1}$ and the recurrence formulas derived above to develop

$$\hat{x}_{j+1} = \hat{x}_j + \hat{p}_j\alpha_j + s_j\omega_j.$$

You may now verify (you don't need to include this with your writeup) that these recursion formulas (with the "hats" removed) will give Algorithm 7.7. BICGSTAB on page 234 of Saad.