Krylov Space Property of Conjugate Gradient Method

Michael Overton, Numerical Computing, Spring 2017

The Conjugate Gradient method is defined on p. 184 of A&G, and the main theorem about the method is on p. 187. The proof of this theorem is not easy, but it can be found in many books, including Trefethen and Bau (however the notation is different, and the assumption that $x_0 = 0$ is made there for convenience). Here, we prove something quite simple, which is still an important basic property of the method. Recall that $\mathrm{span}(v_1, \ldots, v_m)$ means the linear span of vectors $v_1, \ldots, v_n$, that is:

$$\mathrm{span}(v_1, \ldots, v_m) = \{w : w = \gamma_1 v_1 + \cdots + \gamma_m v_m \text{ for some } \gamma_1, \ldots, \gamma_m \in \mathbb{R}\}.$$

Define the kth Krylov space of A with respect to b as

$$\mathcal{K}_k = \mathrm{span}(b, Ab, A^2b, \ldots, A^{k-1}b).$$

Exercise: show that if $w \in \mathcal{K}_k$, then $Aw \in \mathcal{K}_{k+1}$.

Theorem. Assume $x_0 = 0$. Then, for $k = 1, 2, \ldots$, the following statements hold:

1. The $(k-1)$th residual vector r_{k-1} is in $\mathcal{K}_k$
2. The $(k-1)$th direction vector p_{k-1} is in $\mathcal{K}_k$
3. The kth solution approximation vector x_k is in $\mathcal{K}_k$.

The proof is by induction. For $k = 0$, we have by definition that $p_0 = r_0 = b$ and $x_1 = 0 + \alpha_0 p_0$, so, since α_0 is a real scalar, the result holds.

Now assume the inductive hypothesis, namely, these three properties hold for given k; we must show they also hold when k is replaced by $k+1$. We have

$$r_k = r_{k-1} - \alpha_{k-1} A p_{k-1}$$

so, since r_{k-1} and p_{k-1} are both in $\mathcal{K}_k$ by the inductive hypothesis, and A times any vector in $\mathcal{K}_k$ is in $\mathcal{K}_{k+1}$, we have that $r_k \in \mathcal{K}_{k+1}$. Furthermore, we have

$$p_k = r_k + \frac{\delta_k}{\delta_{k-1}} p_{k-1}$$

so, since $p_{k-1} \in \mathcal{K}_k$, and $r_k \in \mathcal{K}_{k+1}$, then $p_k \in \mathcal{K}_{k+1}$. Finally, we have

$$x_{k+1} = x_k + \alpha_k p_k$$

so, since $x_k \in \mathcal{K}_k$, and $p_k \in \mathcal{K}_{k+1}$, then $x_k \in \mathcal{K}_{k+1}$. This proves the theorem.