

Conjugate Gradient is Lanczos in Disguise

Tyler Chen

It's perhaps not so surprising that the Conjugate Gradient and Lanczos algorithms are closely related. After all, they are both Krylov subspace methods for symmetric matrices.

More precisely, the Lanczos algorithm will produce an orthonormal basis for $\mathcal{K}_k(A, b)$, $k = 0, 1, \dots$ if we initialize with initial vector $r_0 = b$. We also know that the Conjugate Gradient residuals form an orthogonal basis for these spaces, which means that the Lanczos vectors must be scaled versions of the Conjugate Gradient residuals.

This relationship provides a way of transferring research about the Lanczos algorithm to be to CG, and visa versa. In fact, the analysis of finite precision CG done by Greenbaum requires viewing CG in terms of the Lanczos recurrence.

In case you're just looking for the punchline, the Lanczos coefficients and vectors can be obtained from the Conjugate Gradient algorithm by the following relationship,

$$q_{j+1} \equiv (-1)^j \frac{r_j}{\|r_j\|}, \quad \beta_j \equiv \frac{\|r_j\|}{\|r_{j-1}\|} \frac{1}{a_{j-1}}, \quad \alpha_j \equiv \left(\frac{1}{a_{j-1}} + \frac{b_j}{a_{j-2}} \right)$$

Note that the indices are offset, because the Lanczos algorithm is started with initial vector q_1 .

Derivation

The derivation of the above result is so much difficult as it is tedious. Before you read my derivation, I would highly recommend trying to derive it on your own, since it's a good chance to improve your familiarity with both algorithms. While I hope that my derivation is not too hard to follow, there are definitely other ways to arrive at the same result, and often to really start to understand something you have to work it out on your own.

Recall that the three term Lanczos recurrence is,

$$Aq_j = \alpha_j q_j + \beta_{j-1} q_{j-1} + \beta_j q_{j+1}$$

In each iteration of CG we update,

$$r_j = r_{j-1} - a_{j-1}Ap_{j-1}, \quad p_j = r_j + b_jp_{j-1}$$

Thus, substituting the expression for p_{j-1} we find,

$$\begin{aligned} r_j &= r_{j-1} - a_{j-1}A(r_{j-1} + b_{j-1}p_{j-2}) \\ &= r_{j-1} - a_{j-1}Ar_{j-1} - a_{j-1}b_{j-1}Ap_{j-2} \end{aligned}$$

Now, rearranging our equation for r_{j-1} we have that $Ap_{j-2} = (r_{j-2} - r_{j-1})/a_{j-2}$. Therefore,

$$r_j = r_{j-1} - a_{j-1}Ar_{j-1} - \frac{a_{j-1}b_{j-1}}{a_{j-2}}(r_{j-2} - r_{j-1})$$

At this point we've found a three term recurrence for r_j , which is a hopeful sign that we are on the right track. We know that the Lanczos vectors are orthonormal and that the recurrence is symmetric, so we'll keep massaging our CG relation to try to get it into that form.

First, let's rearrange terms and regroup them so that the indices and matrix multiply match up with the Lanczos recurrence. This gives,

$$Ar_{j-1} = \left(\frac{1}{a_{j-1}} + \frac{b_{j-1}}{a_{j-2}} \right) r_{j-1} - \frac{b_{j-1}}{a_{j-2}} r_{j-2} - \frac{1}{a_{j-1}} r_j$$

Now, we normalize our residuals as $z_j = r_{j-1}/\|r_{j-1}\|$ so that $r_{j-1} = \|r_{j-1}\|z_j$. Plugging these in gives,

$$\|r_{j-1}\|Az_j = \|r_{j-1}\| \left(\frac{1}{a_{j-1}} + \frac{b_{j-1}}{a_{j-2}} \right) z_j - \|r_{j-2}\| \frac{b_{j-1}}{a_{j-2}} z_j - \|r_j\| \frac{1}{a_{j-1}} z_{j+1}$$

Thus, dividing through by $\|r_{j-1}\|$ we have,

$$Az_j = \left(\frac{1}{a_{j-1}} + \frac{b_{j-1}}{a_{j-2}} \right) z_j - \frac{\|r_{j-2}\|}{\|r_{j-1}\|} \frac{b_{j-1}}{a_{j-2}} z_{j-1} - \frac{\|r_j\|}{\|r_{j-1}\|} \frac{1}{a_{j-1}} z_{j+1}$$

This looks close, but the coefficients for the last two terms have the same formula in the Lanczos recurrence. However, recall that $b_j = \langle r_j, r_j \rangle / \langle r_{j-1}, r_{j-1} \rangle = \|r_j\|^2 / \|r_{j-1}\|^2$. Thus,

$$Az_j = \left(\frac{1}{a_{j-1}} - \frac{b_{j-1}}{a_{j-2}} \right) z_j - \frac{\|r_{j-1}\|}{\|r_{j-2}\|} \frac{1}{a_{j-2}} z_{j-1} - \frac{\|r_j\|}{\|r_{j-1}\|} \frac{1}{a_{j-1}} z_{j+1}$$

We're almost there! While we have the correct for the recurrence, the coefficients from the Lanczos method are always positive. This means that our z_j have the wrong signs. Fixing this gives the relationship,

$$q_{j+1} \equiv (-1)^j \frac{r_j}{\|r_j\|}, \quad \beta_j \equiv \frac{\|r_j\|}{\|r_{j-1}\|} \frac{1}{a_{j-1}}, \quad \alpha_j \equiv \left(\frac{1}{a_{j-1}} + \frac{b_j}{a_{j-2}} \right)$$