Resolution of Iterative Inversion Methods Using Partial Reorthogonalization

James G. Berryman

Notes

Suppose we want to invert a square, $n \times n$ symmetric matrix $\mathbf{A}$ and we want to do so using some iterative method like Lanczos, LSQR, conjugate gradients, etc. Then the iterative algorithm generally may be expressed in the form

$$\mathbf{A}\mathbf{Z}_k = \mathbf{Z}_k\mathbf{T}_k + N_{k+1}\mathbf{z}_{k+1}\mathbf{e}_k^T \tag{1}$$

for $k \leq r$ where r is the rank of the matrix $\mathbf{A}$ and the matrix $\mathbf{Z}_k$, given by

$$\mathbf{Z}_k = \begin{pmatrix} \mathbf{z}_1 & \mathbf{z}_2 & \dots & \mathbf{z}_k \end{pmatrix}, \tag{2}$$

is composed of the column vectors $\mathbf{z}_i$ for $i = 1, \dots, k$ that have been generated so far in the iterative process. The other terms in (1) are the unit vector $\mathbf{e}_k$, with its single component in the kth position, and a norming constant N_{k+1}. The formula (1) with $\mathbf{T}$ being a tridiagonal matrix is just the iterative scheme of Lanczos.

Now the model resolution matrix $\mathcal{R}_{model}$ for the iterative scheme at the kth iteration is given by

$$\mathcal{R}_k = \mathbf{Z}_k\mathbf{Z}_k^T. \tag{3}$$

However, care must be taken to make sure that the $\mathbf{z}_i$s are orthogonal as they are expected to be in this scheme. The tridiagonalization process produces a sequence of orthogonal vectors in principle, but in practice the orthogonality may break down after several iterations when the calculations are performed at finite precision. To demonstrate how the lack of orthogonality affects the process consider the following facts. Let the eigenvalues of the matrix $\mathbf{A}$ be ordered so that $\lambda_i^2 \neq 0$ for $i = 1, \dots, r$ and $\lambda_i^2 = 0$ for $i = r+1, \dots, n$. (We write the eigenvalues as squares because $\mathbf{A}$ often takes the form of a normal matrix $\mathbf{A} = \mathbf{M}^T\mathbf{M}$, in which case the eigenvalues of $\mathbf{M}$ are the λ_is.) Then the trace of the matrix is just

$$\mathrm{Tr}\mathbf{A} = \sum_{i=1}^{r} \lambda_i^2. \tag{4}$$

The model resolution $\mathcal{R}_k$ is a projection operator onto a k-dimensional Krylov subspace (the one explored so far by the iterative method) of the n-dimensional space that is both the range and domain of $\mathbf{A}$. Taking the trace of $\mathcal{R}_k\mathbf{A}$ shows that

$$\mathrm{Tr}\left[\mathbf{Z}_k\mathbf{Z}_k^T\mathbf{A}\right] = \mathrm{Tr}\left[\mathbf{Z}_k^T\mathbf{A}\mathbf{Z}_k\right] = \sum_{i=1}^{k} \mathbf{z}_i^T\mathbf{A}\mathbf{z}_i. \tag{5}$$

Let $\mathbf{v}_i$ for $i = 1, \dots, r$ be the normalized eigenvectors associated with the eigenvalues λ_i^2. Then, assuming only that $\mathbf{z}_1$ has no components in the null-space of $\mathbf{A}$, each of the iteration vectors $\mathbf{z}_i$ can be expanded in terms of these eigenvectors with the result that

$$\mathbf{z}_i = \sum_{j=1}^{r} \zeta_{ij}\mathbf{v}_j. \tag{6}$$

Similarly, the eigenvectors can be expanded in terms of the full set of normalized iteration vectors $\mathbf{z}_i$ according to

$$\mathbf{v}_j = \sum_{i=1}^{r} \zeta_{ij} \mathbf{z}_i, \tag{7}$$

where the same set of coefficients ζ_{ij} is used in both expansions and these coefficients must satisfy

$$\sum_{i=1}^{r} \zeta_{ij}^2 = 1 = \sum_{j=1}^{r} \zeta_{ij}^2 \tag{8}$$

in order for both sets of vectors to be normalized to unity.

Substituting (6) into (5), we find easily that

$$\mathrm{Tr}\left[\mathbf{Z}_k \mathbf{Z}_k^T \mathbf{A}\right] = \sum_{j=1}^{r} \lambda_j^2 \left[\sum_{i=1}^{k} \zeta_{ij}^2\right] \leq \sum_{j=1}^{r} \lambda_j^2. \tag{9}$$

The inequality follows, since the coefficients ζ_{ij} are all real, and therefore

$$\sum_{j=1}^{k} \zeta_{ij}^2 \leq \sum_{j=1}^{r} \zeta_{ij}^2 \equiv 1. \tag{10}$$

Thus, we can expect that as long as the vectors returned by the iterative scheme are orthogonal, the effective trace of the operator $\mathbf{T}$ must remain smaller than that of the original operator $\mathbf{A}$. If this constraint is violated, then we know that the set of $\mathbf{z}_i$s generated so far are not orthogonal and furthermore that this lack of orthogonality is having serious negative consequences on our ability to compute the resolution operator correctly.

The examples presented in the talk demonstrated this result with some clarity.