The distance of an eigenvector to a Krylov subspace and the convergence of the Arnoldi method for eigenvalue problems

M. BELLALIJ, G. MEURANT and H. SADOK

May 2016

M. Bellalij, G. Meurant and H. Sadok,

The distance of an eigenvector to a Krylov subspace and the convergence of the Arnoldi method for eigenvalue problems

to appear in Linear Algebra and its Applications

We consider the eigenvalue problem

$$A\,x = \lambda x$$

The Krylov subspace and matrix are defined as

$$\mathcal{K}_k(A, v) = \mathrm{span}\{v, A\,v, \ldots, A^{k-1}v\}$$

$$K_k = \begin{pmatrix} v & Av & \ldots & A^{k-1}v \end{pmatrix}$$

The Arnoldi algorithm

The residual vectors are orthogonal to the Krylov subspace

$$(A\,x^{(k)} - \lambda^{(k)}\,x^{(k)}\,, A^i v) = 0 \qquad \text{for} \quad i = 0, \dots, k-1$$

The orthonormal basis of $\mathcal{K}_k(A, v)$ is computed using

- $AV_k = V_k H_k + h_{k+1,k} v_{k+1} e_k{}^T, \quad v_1 = v/\|v\|$
- $H_k = V_k^* A V_k$

The approximate eigenvalue problem can now be written as

$$V_k^*\,AV_k\,y^{(k)} = H_k y^{(k)} = \lambda^{(k)}\,y^{(k)}$$

which is equivalent to

$$V_k^*(A - \lambda^{(k)} I) V_k y^{(k)} = 0$$

Approximate eigenvalues (the so-called Ritz values) are the eigenvalues of H_k which is upper Hessenberg

Let $\mathcal{P}_k$ be the orthogonal projector onto $\mathcal{K}_k(A, v)$

It can be written as $K_k(K_k^* K_k)^{-1} K_k^*$ or $V_k V_k^*$ depending on the basis we consider

Then, the approximate eigenvalue problem amounts to solving

$$\mathcal{P}_k(Ax - \lambda x) = 0, \quad x \in \mathcal{K}_k(A, v)$$

or in operator form

$$\mathcal{P}_k A \mathcal{P}_k x = \lambda x$$

Let us define $A_k \equiv \mathcal{P}_k A \mathcal{P}_k$. Note that $A_k = V_k H_k V_k^*$

Convergence analysis

Theorem

Let x be an eigenvector of A associated with the eigenvalue λ and
$\gamma_k = \|\mathcal{P}_k A(I - \mathcal{P}_k)\|$
Then the residual norms of the pairs $(\lambda, \mathcal{P}_k x)$ and (λ, x) for the linear operator A_k satisfy

$$\|(A_k - \lambda I)\mathcal{P}_k x\| \leq \gamma_k \|(I - \mathcal{P}_k)x\|$$

$$\|(A_k - \lambda I)x\| \leq \sqrt{|\lambda|^2 + \gamma_k^2}\, \|(I - \mathcal{P}_k)x\|$$

Y. Saad, Numerical methods for large eigenvalue problems, Halstead Press, (1992). Revised edition, SIAM (2011)

Note that $\gamma_k \leq \|A\|$

Therefore, the coefficients on the right-hand sides of these inequalities are at most of the order of $\|(I - \mathcal{P}_k)x\|$

The Theorem states how accurate the *exact eigenpair* is with respect to the *approximate problem*. This is stated in terms of the distance of the exact eigenvector x to the Krylov subspace.

The remaining issue is to compute or to estimate $\|(I - \mathcal{P}_k)x\|$

Upper bounds for the norm have been obtained in

M. Bellalij, Y. Saad and H. Sadok, *Further analysis of the Arnoldi process for eigenvalue problems*, SIAM J. Numer. Anal., v 48 n 2 (2010), pp. 393–407

using several tools: the eigenvectors, the Schur vectors and an approximation theory viewpoint

We remark that $\|(I - \mathcal{P}_k)x\|$ is also involved in the analysis of convergence of harmonic Ritz values and harmonic and refined Ritz vectors

The minimum distance to a subspace

We consider the minimum distance $\min_{x \in \mathcal{X}} \|w - x\|$ of a vector w to $\mathcal{X}$ which is an arbitrary subspace of dimension k ($w \notin \mathcal{X}$)

Given *any* matrix W whose columns give a basis of the subspace $\mathcal{X}$, $x \in \mathcal{X}$ can be written as Wy, where $y \in \mathbb{C}^k$. Hence,

$$\|w - x\|^2 = \|w - Wy\|^2 = w^* w - w^* Wy - y^* W^* w + y^* W^* Wy$$

$$\|w - x\|^2 = \begin{pmatrix} 1 \\ -y \end{pmatrix}^* \underbrace{\begin{pmatrix} w^* w & w^* W \\ W^* w & W^* W \end{pmatrix}}_{\equiv C} \begin{pmatrix} 1 \\ -y \end{pmatrix}$$

$$\min_{x \in \mathcal{X}} \|w - x\|^2 = \min_{y \in \mathbb{C}^k} \|w - Wy\|^2 = \min_{y \in \mathbb{C}^k} \begin{pmatrix} 1 \\ y \end{pmatrix}^* C \begin{pmatrix} 1 \\ y \end{pmatrix}$$

Lemma

Let $\mathcal{X}$ be an arbitrary subspace of dimension k in $\mathbb{C}^N$ with a basis $W = [w_1, \cdots, w_k]$ and let $w \notin \mathcal{X}$. Let $\mathcal{P}$ be the orthogonal projector onto $\mathcal{X}$. Then,

$$\|(I - \mathcal{P})w\|^2 = \frac{1}{e_1^T C^{-1} e_1}$$

where

$$C = \begin{pmatrix} w^* w & w^* W \\ W^* w & W^* W \end{pmatrix}$$

We have

$$\|(I - \mathcal{P})w\|^2 = w^* w - w^* \mathcal{P} w = w^* w - w^* W (W^* W)^{-1} W^* w$$

The right-hand side is the Schur complement of the (1,1) entry of C which is the inverse of the (1,1) entry of C^{-1}

The minimum distance of an eigenvector to a Krylov subspace

Hypothesis H:

Let $k < N$. The matrix A is diagonalizable as $A = X\Lambda X^{-1}$ where X is the matrix of the normalized eigenvectors and Λ is the diagonal matrix of the eigenvalues denoted as λ_i, $i = 1, \ldots, N$

We assume that there are at least $k+1$ distinct eigenvalues (which we number from 1 to $k+1$) and that the first $k+1$ components of $\alpha = X^{-1}v$ are different from zero

We are interested in the convergence to a given simple eigenvalue which is indexed by 1, that is, we consider x_1, the first column of X corresponding to λ_1. We have assumed that $[X^{-1}v]_1 \neq 0$ and $\|x_j\| = 1$ for all j

We would like to compute $\|(I - \mathcal{P}_k)x_1\|$ where $\mathcal{P}_k$ is the orthogonal projector onto $\mathcal{K}_k(A, v)$

Noting that $C = [w, W]^*[w, W]$ and applying the previous lemma we obtain

Corollary

Let A be a general square matrix, $x_1 \notin \mathcal{K}_k(A, v)$ be an eigenvector of A and L_{k+1} be the rectangular matrix of $\mathbb{C}^{N\times(k+1)}$

$$L_{k+1} = \begin{pmatrix} \alpha_1\, x_1, & v, & A\, v, & \dots, & A^{k-1}\, v \end{pmatrix}$$

We assume that $\alpha \in \mathbb{C}$, $\alpha_1 \neq 0$ and L_{k+1} is of rank $k+1$. Then

$$\|(I - \mathcal{P}_k)\, \alpha_1\, x_1\|^2 = \frac{1}{e_1^T\, (L_{k+1}^* L_{k+1})^{-1}\, e_1}$$

Theorem

Let A be a diagonalizable matrix satisfying hypothesis H, α a vector of $\mathbb{C}^k$ with components α_j such that the starting vector is $v = X\alpha$ and D_α be a diagonal matrix with $(D_\alpha)_{i,i} = \alpha_i$ and

$$W_{k+1} = \begin{pmatrix} 1 & 1 & \lambda_1 & \dots & \lambda_1^{k-1} \\ 0 & 1 & \lambda_2 & \dots & \lambda_2^{k-1} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 1 & \lambda_i & \dots & \lambda_i^{k-1} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 1 & \lambda_N & \dots & \lambda_N^{k-1} \end{pmatrix}$$

Let $\tilde{M}_{k+1} = L_{k+1}^* L_{k+1}$. Then

$$\|(I - \mathcal{P}_k)\,\alpha_1\, x_1\|^2 = \frac{1}{(\tilde{M}_{k+1}^{-1})_{1,1}} = \frac{1}{e_1^T\,(W_{k+1}^*\, D_\alpha^* (X^* X) D_\alpha W_{k+1})^{-1}\, e_1}$$

When A is normal the preceding formula simplifies to

$$\|(I - \mathcal{P}_k)\,\alpha_1\, x_1\|^2 = \frac{1}{e_1^T\,(W_{k+1}^*\, D_\alpha^* D_\alpha W_{k+1})^{-1}\, e_1}$$

The proof is obtained by factoring L_{k+1}

$$
\begin{aligned}
L_{k+1} &= [\alpha_1 x_1, v, A\,v, \ldots, A^{k-1}\,v] \\
&= X\,[\alpha_1\,e_1, \alpha, \Lambda\,\alpha, \ldots, \Lambda^{k-1}\,\alpha] \\
&= X\,D_\alpha\,W_{k+1}
\end{aligned}
$$

Now, using Cramer's rule, we have

$$
(\tilde{M}_{k+1}^{-1})_{1,1} = \frac{\det(\hat{M}_{k+1})}{\det(\tilde{M}_{k+1})}
$$

where $\hat{M}_{k+1}$ is equal to $\tilde{M}_{k+1}$ except for the first column which is replaced by the first column of the identity matrix

Obviously $\det(\tilde{M}_{k+1}) = \det((\tilde{M}_{k+1})_{[2:k+1],[2:k+1]})$ and

$$
\begin{aligned}
(\tilde{M}_{k+1})_{[2:k+1],[2:k+1]} &= \begin{pmatrix} 0 & \\ \vdots & I_k \\ 0 & \end{pmatrix} \tilde{M}_{k+1} \begin{pmatrix} 0 & \cdots & 0 \\ & I_k & \end{pmatrix} \\
&= \begin{pmatrix} 0 & \\ \vdots & I_k \\ 0 & \end{pmatrix} W_{k+1}^* D_\alpha^* (X^* X) D_\alpha W_{k+1} \begin{pmatrix} 0 \\ I_k \end{pmatrix} \\
&= \mathcal{V}_k^* D_\alpha^* (X^* X) D_\alpha \mathcal{V}_k
\end{aligned}
$$

with

$$
\mathcal{V}_k = \begin{pmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{k-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{k-1} \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_N & \cdots & \lambda_N^{k-1} \end{pmatrix}
$$

But

$$
\mathcal{V}_k^* D_\alpha^* (X^* X) D_\alpha \mathcal{V}_k = M_k \equiv K_k^* K_k
$$

Lemma (J. Duintjer Tebbens and G.M.)

Let A *be a diagonalizable matrix with a spectral decomposition* $A = X\Lambda X^{-1}$ *Then*

$$\det(M_k) = \sum_{I_k} \left| \sum_{J_k} \det(X_{I_k,J_k}) \alpha_{j_1} \cdots \alpha_{j_k} \prod_{\substack{j_1 \le j_p < j_q \le j_k \\ j_p, j_q \in J_k}} (\lambda_{j_q} - \lambda_{j_p}) \right|^2$$

where the summations are over all sets of indices I_k *and* J_k *defined as* I_ℓ *to be a set of* ℓ *indices* $(i_1, \quad i_2, \quad \ldots, \quad i_\ell)$ *such that* $1 \le i_1 < \cdots < i_\ell \le N$ *and* J_ℓ *is similar with* i *replaced by* j*,* X_{I_ℓ,J_ℓ} *is the submatrix of* X *whose row and column indices of entries are defined respectively by* I_ℓ *and* J_ℓ *and* $\alpha = X^{-1}v$*. The product is on all pairs of indices that belong to* J_k *(the product being equal to 1 if* $k = 1$*)*

If the matrix A *is normal then with* $\alpha = X^* v$

$$\det(M_k) = \sum_{I_k} \left[\prod_{j=1}^{k} |\alpha_{i_j}|^2 \right] \prod_{\substack{i_1 \le i_p < i_q \le i_k \\ i_p, iq \in I_k}} |\lambda_{i_q} - \lambda_{i_p}|^2$$

Lemma

Let A be a diagonalizable matrix with a spectral decomposition $A = X\Lambda X^{-1}$. Then

$$\det(\tilde{M}_{k+1}) = |\alpha_1|^2 \sum_{I_{k+1}} \left| \sum_{\hat{J}_{k+1}} \det(X_{I_{k+1},\hat{J}_{k+1}})\, \alpha_{j_2} \cdots \alpha_{j_{k+1}} \prod_{\substack{1<j_2\le j_p<j_q\le j_{k+1}\\ j_p,j_q\in\hat{J}_{k+1}}} (\lambda_{j_q}-\lambda_{j_p}) \right|^2$$

where the summation with $\hat{J}_{k+1}$ is over all sets of indices $\{1, j_2, \ldots, j_{k+1}\}$ such that $1 < j_2 < \cdots < j_{k+1} \le N$
If the matrix A is normal then

$$\det(\tilde{M}_{k+1}) = |\alpha_1|^2 \sum_{\hat{I}_{k+1}} \left[\prod_{j=2}^{k+1} |\alpha_{i_j}|^2 \right] \prod_{\substack{1<i_2\le i_p<i_q\le i_{k+1}\\ i_p,i_q\in\hat{I}_{k+1}}} |\lambda_{i_q}-\lambda_{i_p}|^2$$

and the summation with $\hat{I}_{k+1}$ is over all sets of indices $\{i_2, \ldots, i_{k+1}\}$ such that $1 < i_2 < \cdots < i_{k+1} \le N$

Let $G = XD_\alpha W_{k+1}$, then

$$\tilde{M}_{k+1} = G^* G$$

We use the Cauchy-Binet formula

$$\det(\tilde{M}_{k+1}) = \sum_{I_{k+1}} |\det(G_{I_{k+1},:})|^2$$

It yields

$$\det(G_{I_{k+1},:}) = \sum_{J_{k+1}} \det(X_{I_{k+1},J_{k+1}})\, \alpha_{j_1} \cdots \alpha_{j_{k+1}} \det(\mathcal{W}(j_1, \ldots, j_{k+1}))$$

where $\mathcal{W}(j_1, \ldots, j_{k+1})$ is obtained from the rows $j_1, \ldots, j_{k+1}$ of W_{k+1}

$\det(\mathcal{W}(j_1,\ldots,j_{k+1})) = 0$ if $1 \notin \{j_1, j_2, \ldots, j_{k+1}\}$

But the indices in J_{k+1} are strictly ordered, so the determinant is different from zero only if $j_1 = 1$

The sum over the sets J_{k+1} reduces to a sum over sets of indices $\hat{J}_{k+1}$ which are $\{1, j_2, \ldots, j_{k+1}\}$ with $1 < j_2 < \cdots < j_{k+1} \leq N$

$$\det(\mathcal{W}(1, j_2, \ldots, j_{k+1})) = \det \begin{pmatrix} 1 & \lambda_{j_2} & \cdots & \lambda_{j_2}^{k-1} \\ 1 & \lambda_{j_3} & \cdots & \lambda_{j_3}^{k-1} \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_{j_{k+1}} & \cdots & \lambda_{j_{k+1}}^{k-1} \end{pmatrix}$$

It yields

$$\prod_{1<j_2\leq j_p<j_q\leq j_{k+1}} (\lambda_{j_q} - \lambda_{j_p})$$

where the indices j_p, j_q have to belong to the set $\{j_2, j_3, \ldots, j_{k+1}\}$

Main result

Let A *be a diagonalizable matrix satisfying hypothesis H. The distance of the eigenvector* x_1 *to the* Krylov *subspace* $\mathcal{K}_k(A, v)$ *is given by*

$$\|(I - \mathcal{P}_k)\, x_1\|^2 = \frac{N}{D}$$

with

$$N = \sum_{I_{k+1}} \left| \sum_{\hat{J}_{k+1}} \det(X_{I_{k+1},\hat{J}_{k+1}})\, \alpha_{j_2} \cdots \alpha_{j_{k+1}} \prod_{\substack{1<j_2\leq j_p<j_q\leq j_{k+1} \\ j_p, j_q \in \hat{J}_{k+1}}} (\lambda_{j_q} - \lambda_{j_p}) \right|^2$$

$$D = \sum_{I_k} \left| \sum_{J_k} \det(X_{I_k,J_k}) \alpha_{j_1} \cdots \alpha_{j_k} \prod_{\substack{j_1\leq j_p<j_q\leq j_k \\ j_p, j_q \in J_k}} (\lambda_{j_q} - \lambda_{j_p}) \right|^2$$

with $\alpha = X^{-1} v$

If, in addition, A is normal we have

$$N = \sum_{\hat{I}_{k+1}} \left[\prod_{j=2}^{k+1} |\alpha_{i_j}|^2 \right] \prod_{\substack{1 < i_2 \leq i_p < i_q \leq i_{k+1} \\ i_p, i_q \in \hat{I}_{k+1}}} |\lambda_{i_q} - \lambda_{i_p}|^2$$

$$D = \sum_{I_k} \left[\prod_{j=1}^{k} |\alpha_{i_j}|^2 \right] \prod_{\substack{i_1 \leq i_p < i_q \leq i_k \\ i_p, i_q \in I_k}} |\lambda_{i_q} - \lambda_{i_p}|^2$$

with $\alpha = X^* v$

The formula for the non-normal case is difficult to interpret since there is a strong dependence on the eigenvectors through determinants of submatrices of X

The result for normal matrices can be written in a different way. Let

$$I_k = \mathcal{I}_1 \cup \mathcal{I}_k, \quad \mathcal{I}_1 = \{1, i_2, \dots, i_k\}, \quad \mathcal{I}_k = \{i_1, i_2, \dots, i_k, | \, i_1 > 1\}.$$

Let us denote the sums over these two sets of indices by $S_{\mathcal{I}_1}$ and $S_{\mathcal{I}_k}$

$$S_{\mathcal{I}_1} = |\alpha_1|^2 \sum_{\substack{\{i_2,\dots,i_k\} \\ i_2>1}} \left[\prod_{j=2}^{k} |\alpha_{i_j}|^2 \right] \prod_{\substack{1\le i_p<i_q\le i_k \\ i_p,i_q\in\{1,i_2,\dots,i_k\}}} |\lambda_{i_q} - \lambda_{i_p}|^2$$

$$S_{\mathcal{I}_k} = \sum_{\mathcal{I}_k} \left[\prod_{j=1}^{k} |\alpha_{i_j}|^2 \right] \prod_{\substack{1<i_1\le i_p<i_q\le i_k \\ i_p,i_q\in\mathcal{I}_k}} |\lambda_{i_q} - \lambda_{i_p}|^2$$

$S_{\mathcal{I}_k}$ is equal to the numerator N

$$\frac{N}{D} = \frac{S_{\mathcal{I}_k}}{S_{\mathcal{I}_1} + S_{\mathcal{I}_k}} = \frac{1}{1 + \frac{S_{\mathcal{I}_1}}{S_{\mathcal{I}_k}}}$$

as long as $S_{\mathcal{I}_k} \neq 0$

The eigenvalue of interest λ_1 does not appear in $S_{\mathcal{I}_k}$

We see that the distance of x_1 to the Krylov subspace is equal to 1 if and only if $S_{\mathcal{I}_1} = 0$

It is small if $S_{\mathcal{I}_1}/S_{\mathcal{I}_k}$ is large

A small example

We consider a normal matrix A of order 4 with eigenvalues $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ and $k = 2$

The sets of indices in I_3 are $(1, 2, 3)$, $(1, 2, 4)$, $(1, 3, 4)$

Therefore, the sets of indices in $\hat{I}_3$ are $(2, 3)$, $(2, 4)$, $(3, 4)$. The sets of indices in I_2 are

$$(1, 2),\ (1, 3),\ (1, 4),\ (2, 3),\ (2, 4),\ (3, 4)$$

The three first pairs are in $\mathcal{I}_1$ and the four last ones are in $\mathcal{I}_2$

We see that $\mathcal{I}_2$ is identical to $\hat{I}_3$

Let $\beta_i = |\alpha_i|^2$. We have

$$\begin{aligned} S_{\mathcal{I}_1} &= \beta_1 \left[\beta_2|\lambda_2 - \lambda_1|^2 + \beta_3|\lambda_3 - \lambda_1|^2 + \beta_4|\lambda_4 - \lambda_1|^2\right] \\ S_{\mathcal{I}_2} &= \beta_2\beta_3|\lambda_3 - \lambda_2|^2 + \beta_2\beta_4|\lambda_4 - \lambda_2|^2 + \beta_3\beta_4|\lambda_4 - \lambda_3|^2 \end{aligned}$$

The term $S_{\mathcal{I}_1}$ will be large if at least one of the other eigenvalues is "far" from λ_1 and the projection of v on the corresponding eigenvector is not too small

The other term $S_{\mathcal{I}_2}$ is small if the products of the pairwise distances between the other eigenvalues with the moduli of the projections of v are small

If only one of the terms in the sum is not small, $S_{\mathcal{I}_2}$ cannot be small

Let us assume that we have a cluster of three distinct eigenvalues $\lambda_2, \lambda_3, \lambda_4$ whose small pairwise distances are of order ε and another complex eigenvalue λ_1 whose pairwise distances to the three other ones are of order 1

Assume also that all the β_i's are non zero and that no one is very small

Then $S_{\mathcal{I}_2} = c\varepsilon^2$ where $c \gg \varepsilon$ and the ratio N/D is equal to

$$\frac{1}{1 + \frac{S_{\mathcal{I}_1}}{c\varepsilon^2}} = \frac{c\varepsilon^2}{c\varepsilon^2 + S_{\mathcal{I}_1}} = \mathcal{O}(\varepsilon^2)$$

since $S_{\mathcal{I}_1}$ is of order 1

We see that with this distribution of eigenvalues $\|(I - \mathcal{P}_2)\,x_1\|$ is small of order ε

Let us now assume that A is real. Then the eigenvalues arise as real numbers or complex conjugate pairs

Let λ_1 be a complex eigenvalue with $\lambda_2 = \overline{\lambda_1}$

Assume that λ_3 and λ_4 are complex conjugate or real with $|\lambda_4 - \lambda_3| = \varepsilon$ and the distances to λ_1 and λ_2 are of order 1

Then, $S_{\mathcal{I}_2}$ is not small unless β_2 is small. Generally, the ratio N/D is not small

Let us consider the normal matrix

$$A = \begin{pmatrix} 1.894 & 0.09975 & -0.2124 & -0.3811 \\ 0.4172 & 1.137 & 0.3024 & 0.8461 \\ 0.1531 & -0.7204 & 1.649 & -0.1911 \\ 0.05344 & -0.6726 & -0.6651 & 1.32 \end{pmatrix}$$

and

$$v = \begin{pmatrix} -0.3019 \\ -0.2286 \\ -0.8316 \\ 0.4063 \end{pmatrix}$$

The eigenvalues of A are

$$\left(1+i, \quad 1-i, \quad 2-0.01i, \quad 2+0.01i\right)$$

The distance is $\|x_1 - \mathcal{P}_2 x_1\| = 0.67682$

The two sums are

$$S_{\mathcal{I}_1} = 0.15514, \quad S_{\mathcal{I}_2} = 0.13115$$

$S_{\mathcal{I}_1}$ is not large and $S_{\mathcal{I}_2}$ is not small

The two Ritz values at iteration 2 are real, being $(1.8123, 1.1039)$

$$\|(A_2-(1+i)I)P_2x_1\| = 0.67421, \gamma_2 = 0.99644, \|x_1-\mathcal{P}_2x_1\| = 0.67682$$

$$\gamma_2\,\|x_1 - \mathcal{P}_2x_1\| = 0.67441$$

and

$$\|(A_2 - (1+i)I)x_1\| = 1.1708, \quad \sqrt{1+\gamma_2^2} = 1.7300$$

$$\sqrt{1+\gamma_2^2}\,\|x_1 - \mathcal{P}_2x_1\| = 1.1709$$

The two bounds are close to the values of the residual norms

Conclusion

- We exhibited an exact expression for the distance of an eigenvector to a Krylov subspace
- It depends in an intricate way on the eigenvalues, the eigenvectors and the starting vector
- The expression is simpler when the matrix is normal
- This gives some bounds for residual norms in the Arnoldi algorithm

Finally, let us remember the happy days. . .

Marrakech, 2011