KRYLOV SPACES, GMRES, SINGULAR VALUES AND EQUATIONS OF THE SECOND KIND

30th May 2007

1 Krylov spaces

Notations. Let H be a complex Hilbert space and $A : H \to H$ be an invertible bounded linear operator. Given $f \neq 0$ we consider the equation

$$Au = f.$$

Krylov subspaces are defined as

$$\mathcal{W}^{(n)} := \mathbb{C}\langle f, Af, \dots, A^{n-1}f\rangle.$$

Notice that for all n

$$\mathcal{W}^{(n)} \subseteq \mathcal{W}^{(n+1)}$$

and

$$A\mathcal{W}^{(n)} = \mathbb{C}\langle Af, A^2 f, \dots, A^n f\rangle \subseteq \mathcal{W}^{(n+1)}.$$

Proposition *If* $\mathcal{W}^{(n)} = \mathcal{W}^{(n+1)}$ *then*

$$\mathcal{W}^{(n)} = \mathcal{W}^{(n+1)} = \mathcal{W}^{(n+2)} = \dots$$

Proof. It is obvious that

$$\mathcal{W}^{(n+1)} = \mathcal{W}^{(n)} + \mathbb{C}\langle A^n f\rangle.$$

Therefore

$$\mathcal{W}^{(n+1)} = \mathcal{W}^{(n)} \qquad \Longleftrightarrow \qquad A^n f \in \mathcal{W}^{(n)}.$$

However, if $A^n f \in \mathcal{W}^{(n)}$, then $A^{n+1} f \in A\mathcal{W}^{(n)} \subseteq \mathcal{W}^{(n+1)}$, which is equivalent to $\mathcal{W}^{(n+2)} = \mathcal{W}^{(n+1)}$. We then proceed by induction. □

Proposition *If* $A\mathcal{W}^{(n)} = \mathcal{W}^{(n+1)}$, *then*

$$A\mathcal{W}^{(k)} = \mathcal{W}^{(k+1)}, \qquad \forall k \geq n.$$

Proof. Since

$$\mathcal{W}^{(n+1)} = \mathbb{C}\langle f\rangle + A\mathcal{W}^{(n)}$$

it follows that

$$\mathcal{W}^{(n+1)} = A\mathcal{W}^{(n)} \qquad \Longleftrightarrow \qquad f \in A\mathcal{W}^{(n)}.$$

However, $A\mathcal{W}^{(n)} \subseteq A\mathcal{W}^{(k)}$, for all $k \geq n$, and therefore the result is obvious. □

Proposition *For given n*

$$\mathcal{W}^{(n)} = \mathcal{W}^{(n+1)} \qquad \Longleftrightarrow \qquad A\mathcal{W}^{(n)} = \mathcal{W}^{(n+1)}.$$

In this case

$$\begin{aligned} \mathcal{W}^{(n)} &= \mathcal{W}^{(n+1)} = \mathcal{W}^{(n+2)} = \dots \\ &= A\mathcal{W}^{(n)} = A\mathcal{W}^{(n+1)} = \dots \end{aligned}$$

Proof. **(a)** If $\mathcal{W}^{(n)} = \mathcal{W}^{(n+1)}$ we take the minimun value N such that

$$\mathcal{W}^{(N)} = \mathcal{W}^{(N+1)}.$$

Then

$$A^N f = \alpha_0 f + \ldots + \alpha_{N-1} A^{N-1} f.$$

If $\alpha_0 = 0$, we would hace

$$A(\alpha_1 f + \ldots + \alpha_{N-1} A^{N-2} f - A^{N-1} f) = 0$$

and thus

$$A^{N-1} f = \alpha_1 f + \ldots + \alpha_{N-1} A^{N-2} f \in \mathcal{W}^{(N-1)},$$

which implies that $\mathcal{W}^{(N-1)} = \mathcal{W}^{(N)}$, that contradicts our choice of N. Hence $\alpha_0 \neq 0$ and

$$f = -\frac{\alpha_1}{\alpha_0} A f - \ldots - \frac{\alpha_{N-1}}{\alpha_0} A^{N-1} f + A^N f \in A\mathcal{W}^{(N)} \subseteq A\mathcal{W}^{(n)}$$

and thus $\mathcal{W}^{(n+1)} = A\mathcal{W}^{(n)}$.

(b) Assume now that $\mathcal{W}^{(n+1)} = A\mathcal{W}^{(n)}$ and take the minimum value of N such that

$$\mathcal{W}^{(N+1)} = A\mathcal{W}^{(N)}.$$

Then $f \in A\mathcal{W}^{(N)}$ and

$$f = \beta_1 A f + \ldots + \beta_N A^N f$$

with $\beta_N \neq 0$ ($\beta_N = 0$ would give a smaller value of N), so

$$A^N f = \frac{1}{\beta_N} f - \frac{\beta_1}{\beta_N} A f - \ldots - \frac{\beta_{N-1}}{\beta_N} A^{N-1} f \in \mathcal{W}^{(N)}$$

which si equivalent to $\mathcal{W}^{(N)} = \mathcal{W}^{(N+1)}$ and therefore implies that $\mathcal{W}^{(n)} = \mathcal{W}^{(n+1)}$. □

Remark. For a given value n

$$\begin{aligned}\mathcal{W}^{(n)} = \mathcal{W}^{(n+1)} &\iff A^n f \in \mathcal{W}^{(n)} \\ &\iff f \in A\mathcal{W}^{(n)} \\ &\iff A\mathcal{W}^{(n)} = \mathcal{W}^{(n+1)} \\ &\iff A^{-1} f \in \mathcal{W}^{(n)}.\end{aligned}$$

In this case we can take the minimum value N that makes all the properties hold and we have

$$\begin{aligned} A^N f &= \alpha_0 f + \ldots + \alpha_{N-1} A^{N-1} f, \qquad \alpha_0 \neq 0, \\ f &= \beta_1 A f + \ldots + \beta_N A^N f, \qquad \beta_N \neq 0. \end{aligned}$$

Notation. Let

$$N := \sup_n \dim \mathcal{W}^{(n)}.$$

There are two possibilities:

(a) $N = \infty$ means that

$$\mathcal{W}^{(n)} \subsetneq \mathcal{W}^{(n+1)}, \qquad \forall n$$

or equivalently $u = A^{-1} f \notin \mathcal{W}^{(n)}$ for all n.

(b) $N < \infty$ means that $\mathcal{W}^{(N)} = \mathcal{W}^{(N+1)}$ and

$$N = \inf\{n \,|\, \mathcal{W}^{(n)} = \mathcal{W}^{(n+1)}\}$$

Up to this value

$$\dim \mathcal{W}^{(n)} = \dim A\mathcal{W}^{(n)} = n$$

2 The Galerkin–Krylov discretization

Notations. Let

$$P_n : H \to \mathcal{W}^{(n)}$$

be the orthogonal projection onto $\mathcal{W}^{(n)}$ and

$$A_n := P_n A|_{\mathcal{W}^{(n)}} : \mathcal{W}^{(n)} \to \mathcal{W}^{(n)}.$$

Remarks. Notice that since $f \in \mathcal{W}^{(n)}$ for all n, then $P_n f = f$. The equation

$$A_n y_n = f = P_n f$$

can be written equivalently as

$$\left| \begin{array}{l} y_n \in \mathcal{W}^{(n)}, \\ \langle A y_n, w \rangle = \langle f, w \rangle, \qquad \forall w \in \mathcal{W}^{(n)}. \end{array} \right.$$

It looks like a Galerkin discretization, but the sequence of spaces depends on the right–hand side f.

Proposition *If $\mathcal{W}^{(N)} = \mathcal{W}^{(N+1)}$ (i.e. $N < \infty$) then:*

(a) $Aw_N \in \mathcal{W}^{(N)}$ *for all* $w_N \in \mathcal{W}^{(N)}$.

(b) A_N *is invertible.*

(c) $A|_{\mathcal{W}^{(N)}} : \mathcal{W}^{(N)} \to \mathcal{W}^{(N)}$ *is and isomorphism.*

(d) $A^{-1}f \in \mathcal{W}^{(N)}$.

In this case, the unique solution to

$$A_N u_N = f$$

is $u_N = u = A^{-1}f$.

Proof. The hypothesis implies that $A\mathcal{W}^{(N)} = \mathcal{W}^{(N)}$ and we thus have (a) and $A_N = A|_{\mathcal{W}^{(N)}}$. Therefore $\ker A_N \subseteq \ker A = 0$. Since $\mathcal{W}^{(N)}$ is finite dimensional this implies thatA_N is invertible, i.e. (b). The remaining assertions are straightforward. $\square$

3 Arnoldi and GMRES iterations

Arnoldi method. It consists of a modified version of the Gram–Schmidt method to compute an increasing orthonormal system spanning the Krylov spaces:

$$\left| \begin{array}{l} t_1 := \dfrac{1}{\|f\|} f \\ t_{n+1} := \dfrac{1}{\|(I-P_n)At_n\|} (I-P_n)At_n, \qquad n \geq 1. \end{array} \right.$$

If $(I - P_n)At_n = 0$ the method stops.

Proposition *If the method has not stopped at n, then*

$$\mathcal{W}^{(n)} = \mathbb{C}\langle t_1, \ldots, t_n \rangle$$

and $\langle t_i, t_j \rangle = \delta_{ij}$.

Proof. (Induction) If $\mathcal{W}^{(n-1)} = \mathbb{C}\langle t_1, \ldots, t_{n-1} \rangle$, then

$$\begin{array}{rcl} \mathcal{W}^{(n)} & = & \mathbb{C}\langle f \rangle + A\mathcal{W}^{(n-1)} \\ & = & \mathbb{C}\langle f \rangle + \mathbb{C}\langle At_1, \ldots, At_{n-2}, At_{n-1} \rangle \\ & = & \mathbb{C}\langle f \rangle + A\mathcal{W}^{(n-2)} + \mathbb{C}\langle At_{n-1} \rangle \\ & = & \mathcal{W}^{(n-1)} + \mathbb{C}\langle At_{n-1} \rangle. \end{array}$$

Then $(I - P_{n-1})At_{n-1}$ is orthogonal to $\mathcal{W}^{(n-1)}$ and the only remaining step is normalization. $\square$

When the method stops. It is simple to check that

$$\begin{aligned}(I-P_n)At_n = 0 \quad &\Longleftrightarrow \quad P_nAt_n = At_n \\ &\Longleftrightarrow \quad At_n \in \mathcal{W}^{(n)} \\ &\Longleftrightarrow \quad \mathcal{W}^{(n+1)} = \mathcal{W}^{(n)}.\end{aligned}$$

Therefore:

- if $N = \infty$ the Arnoldi iterations never stop;
- if $N < \infty$, we can take
$$t_{N+1} = t_{N+2} = \ldots = 0.$$

Proposition *For all n*

$$\langle At_n, t_{n+1}\rangle = \|(I-P_n)At_n\|.$$

Proof. If $t_{n+1} \neq 0$, then

$$\|(I-P_n)At_n\|\, t_{n+1} = (I-P_n)At_n$$

and therefore multiplying by t_{n+1}

$$\|(I-P_n)At_n\| = \langle At_n, t_{n+1}\rangle - \langle P_nAt_n, t_{n+1}\rangle = \langle At_n, t_{n+1}\rangle,$$

since $t_{n+1} \perp \mathcal{W}^{(n)}$. If $t_{n+1} = 0$, both sides of the equality vanish. □

GMRES iterations. We define u_n as the unique solution of

$$\|f - Au_n\| = \text{min!}, \qquad u_n \in \mathcal{W}^{(n)}.$$

The residual is given by

$$r_n := f - Au_n.$$

Notice that if $\mathcal{W}^{(N)} = \mathcal{W}^{(N+1)}$, then $A^{-1}f = u \in \mathcal{W}^{(N)}$ and $u_N = u$, so the iterations can be stopped. Reciprocally $r_N = 0$ implies that $\mathcal{W}^{(N)} = \mathcal{W}^{(N+1)}$.

Proposition *For all n*

(a) $r_n \in \mathcal{W}^{(n+1)}$.

(b) $\|r_{n-1}\| \leq \|r_n\|$.

(c) $r_n \perp A\mathcal{W}^{(n)}$.

Proof. Since $r_n = f - Au_n \in f + A\mathcal{W}^{(n)} \subset \mathcal{W}^{(n+1)}$ the first result is obvious. On the other hand, since $u_n \in \mathcal{W}^{(n)} \subset \mathcal{W}^{(n+1)}$

$$\|r_{n+1}\| = \|f - Au_{n+1}\| \leq \|f - Au_n\| = \|r_n\|$$

which proves the second assertion. Finally, the third statement is classical: the minimization process is equivalent to

$$\frac{1}{2}(\mathrm{re})\langle Au_n, Au_n\rangle - \mathrm{re}\langle f, Au_n\rangle = \min!, \qquad u_n \in \mathcal{W}^{(n)}$$

itself equivalent to

$$\left| \begin{array}{l} u_n \in \mathcal{W}^{(n)}, \\ \langle Au_n, Av\rangle = \langle f, Av\rangle, \qquad \forall v \in \mathcal{W}^{(n)}, \end{array} \right.$$

or to

$$\left| \begin{array}{l} u_n \in \mathcal{W}^{(n)}, \\ \langle Au_n - f, w\rangle = 0, \qquad \forall w \in A\mathcal{W}^{(n)}. \end{array} \right.$$

□

4 Residual reduction in GMRES

Notation. Let z_n be an orthonormal system such that for all n

$$A\mathcal{W}^{(n)} = \mathbb{C}\langle z_1, \ldots, z_n\rangle.$$

Thus

$$A\mathcal{W}^{(n+1)} = \mathbb{C}\langle z_{n+1}\rangle \oplus A\mathcal{W}^{(n)},$$

with orthogonal sum. If $A\mathcal{W}^{(N)} = A\mathcal{W}^{(N+1)}$ (i.e. $\mathcal{W}^{(N)} = \mathcal{W}^{(N+1)}$) we take $z_{N+1} = z_{N+2} = \ldots = 0$.

Proposition *For all n*

$$r_n - r_{n-1} \in \mathbb{C}\langle z_n\rangle.$$

Proof. We have

$$r_n - r_{n-1} = (f - Au_n) - (f - Au_{n-1}) = A(u_{n-1} - u_n) \in A\mathcal{W}^{(n)}.$$

At the same time

$$\left. \begin{array}{l} r_n \perp A\mathcal{W}^{(n)} \supseteq A\mathcal{W}^{(n-1)} \\ r_{n-1} \perp A\mathcal{W}^{(n-1)} \end{array} \right| \quad \Longrightarrow \quad r_n - r_{n-1} \perp A\mathcal{W}^{(n-1)}.$$

The orthogonal complement of $A\mathcal{W}^{(n-1)}$ in $A\mathcal{W}^{(n)}$ is $\mathbb{C}\langle z_n\rangle$, which finishes the proof. □

Proposition

$$\|r_n\| = \|r_{n-1}\| \qquad \Longleftrightarrow \qquad u_n = u_{n-1}.$$

Proof. Since $r_n - r_{n-1} \in \mathbb{C}\langle z_n \rangle$, then

$$r_{n-1} = r_n + \gamma_n\, z_n.$$

However, $z_n \in A\mathcal{W}^{(n)}$ and $r_n \perp A\mathcal{W}^{(n)}$, from where

$$\|r_{n-1}\|^2 = \|r_n\|^2 + |\gamma_n|^2 \geq \|r_n\|^2.$$

Equality means that $\gamma_n = 0$ and therefore $r_{n-1} = r_n$, which implies $u_{n-1} = u_n$.

Proposition *Assume that $z_n \neq 0$. Then A_n is singular if and only if $P_n z_n = 0$.*

Proof. If A_n is singular, then there exists $0 \neq w_n \in \mathcal{W}^{(n)}$ such that $P_n A w_n = P_n w_n = 0$. Let $v_n := A w_n \in A\mathcal{W}^{(n)}$. Then $P_n v_n = 0$ and $v_n \neq 0$. Notice that

$$v_n \in A\mathcal{W}^{(n)} = \mathbb{C}\langle z_n \rangle \oplus A\mathcal{W}^{(n-1)}$$

(orthogonal sum) and that $A\mathcal{W}^{(n-1)} \subseteq \mathcal{W}^{(n)}$. Hence

$$v_n = \langle v_n, z_n \rangle\, z_n + A w_{n-1}, \qquad w_{n-1} \in \mathcal{W}^{(n-1)}$$

and

$$0 = \langle v_n, z_n \rangle P_n z_n + A w_{n-1}$$

but $\langle z_n, A w_{n-1} \rangle = 0$ and therefore $P_n z_n = 0$ and $v_n \in \mathbb{C}\langle z_n \rangle$.

If $P_n z_n = 0$, since $z_n \in A\mathcal{W}^{(n)}$, then $z_n = A w_n$ with $0 \neq w_n \in \mathcal{W}^{(n)}$. Then

$$0 = P_n z_n = P_n A w_n = A_n w_n.$$

□

Proposition *If A_n is singular, then:*

(a) $\|r_n\| = \|r_{n-1}\|$ *(and $u_n = u_{n-1}$).*

(b) $|\langle t_{n+1}, z_n \rangle| = 1$*, i.e., z_n and t_{n+1} are proportional.*

Proof. Since $r_n - r_{n-1} \in \mathbb{C}\langle z_n \rangle$ and $P_n z_n = 0$, then

$$P_n r_n = P_n r_{n-1} = r_{n-1}$$

($r_{n-1} \in \mathcal{W}^{(n)}$) and therefore

$$\|r_{n-1}\| = \|P_n r_n\| \leq \|r_n\| \leq \|r_{n-1}\|.$$

Since $P_n z_n = 0$ (hence $z_n \perp \mathcal{W}^{(n)}$) and $z_n \in A\mathcal{W}^{(n)} \subseteq \mathcal{W}^{(n+1)}$, then $z_n \in \mathbb{C}\langle t_{n+1} \rangle$. Noticing that both z_n and t_{n+1} hace unit norm, the result follows readily. □

Remark. Proportionality of t_{n+1} and z_n means that $\mathcal{W}^{(n)}$ increases to become $\mathcal{W}^{(n+1)}$ in the same direction as $A\mathcal{W}^{(n-1)}$ increased to become $A\mathcal{W}^{(n)}$.

Proposition (GMRES as a descent method) *If A_n is inversible and y_n is the unique solution to*

$$\left|\begin{array}{l} y_n \in \mathcal{W}^{(n)}, \\ P_n A y_n = f, \end{array}\right.$$

then

$$u_n = \alpha_n y_n + (1-\alpha_n) u_{n-1}, \qquad \alpha_n = \frac{\|r_{n-1}\|^2}{\|f - Ay_n\|^2 + \|r_{n-1}\|^2}.$$

Proof. We begin with a simple remark: $\alpha_n = 1$ if and only if $Ay_n = f$, i.e., $y_n = u = u_n$. Moreover, $Ay_n = f$ if and only if $Ay_n \in \mathcal{W}^{(n)}$. Therefore, eliminating this simple situation we can assume that $\alpha_n \in (0,1)$ and $Ay_n \notin \mathcal{W}^{(n)}$.

Let then $v_n := \alpha_n y_n + (1-\alpha_n) u_{n-1} \in \mathcal{W}^{(n)}$ be defined as above and let

$$\widehat{r}_n := f - Ay_n.$$

To demonstrate that $v_n = u_n$ we have to show that $f - Av_n$ is orthogonal to $A\mathcal{W}^{(n)}$.

Since we have assumed that $Ay_n \notin A\mathcal{W}^{(n-1)} \subset \mathcal{W}^{(n)}$, then

$$\begin{array}{rcl} A\mathcal{W}^{(n)} & = & A\mathcal{W}^{(n-1)} \oplus \mathbb{C}\langle Ay_n \rangle \\ & = & A\mathcal{W}^{(n-1)} \oplus \mathbb{C}\langle A(u_{n-1} - y_n) \rangle \\ & = & A\mathcal{W}^{(n-1)} \oplus \mathbb{C}\langle \widehat{r}_n - r_{n-1} \rangle. \end{array}$$

Notice first that

$$\begin{array}{rcl} f - A\,v_n & = & \alpha_n (f - Ay_n) + (1-\alpha_n)(f - Au_{n-1}) \\ & = & \alpha_n \widehat{r}_n + (1-\alpha_n) r_{n-1}. \end{array}$$

By definition of u_{n-1} and y_n

$$\left.\begin{array}{l} f - Ay_n \perp \mathcal{W}^{(n)} \supseteq A\mathcal{W}^{(n)} \\ f - Au_{n-1} \perp A\mathcal{W}^{(n-1)} \end{array}\right| \quad \Longrightarrow \quad f - Av_n \perp A\mathcal{W}^{(n-1)},$$

so we are left to prove that

$$\alpha_n \widehat{r}_n + (1-\alpha_n) r_{n-1} \perp \widehat{r}_n - r_{n-1},$$

that is (since $\alpha_n = \|r_{n-1}\|^2 / (\|r_{n-1}\|^2 + \|\widehat{r}_n\|^2)$),

$$\|r_{n-1}\|^2\, \widehat{r}_n + \|\widehat{r}_n\|^2\, r_{n-1} \perp \widehat{r}_n - r_{n-1},$$

but

$$\langle \widehat{r}_n, r_{n-1} \rangle = \langle f - Ay_n, r_{n-1} \rangle = 0$$

($r_{n-1} \in \mathcal{W}^{(n)}$), which proves the result. $\square$

Proposition *If A_n is regular and $A_n y_n = f$, then*

$$\|r_n\|^2 = \frac{\|r_{n-1}\|^2 \, \|\widehat{r}_n\|^2}{\|r_{n-1}\|^2 + \|\widehat{r}_n\|^2}, \qquad \widehat{r}_n := f - Ay_n.$$

Then:

- *either $\widehat{r}_n = 0$ and we have reached the exact solution*
- *or $0 < \|r_n\| < \|r_{n-1}\|$.*

Proof. Since $r_n = \alpha_n \widehat{r}_n + (1-\alpha_n) r_{n-1}$ and $\langle \widehat{r}_n, r_{n-1}\rangle = 0$ (see the preceding proof), then

$$\|r_n\|^2 = |\alpha_n|^2 \, \|\widehat{r}_n\|^2 + |1-\alpha_n|^2 \, \|r_{n-1}\|^2$$

and the result follows readily. □

Possibilities in GMRES. Several situations can arise when one does a GMRES iteration (notice that in practice, in finite dimension, one never computes u_n):

(a) $r_{n-1} = 0$; the method has already stopped; A_n is regular and the Krylov spaces are not growing any longer.

(b) $\|r_{n-1}\| = \|r_n\| \neq 0$; in this case $u_n = u_{n-1} \neq u$, so the method has temporarily stuck; necessarily A_n is singular (but Krylov spaces are still growing).

(c) $\|r_n\| < \|r_{n-1}\|$ and necessarily A_n is regular. Then

$$\|r_n\| = \frac{\|r_{n-1}\| \, \|\widehat{r}_n\|}{\sqrt{\|r_{n-1}\|^2 + \|\widehat{r}_n\|^2}}.$$

In particular, it can happen that $\alpha_n = 1$, i.e. $\widehat{r}_n = f - Ay_n = 0$ and $u_n = y_n = u$ is the exact solution. This is the only way to get to the exact solution in a finite number of iterations.

Proposition *In all situations*

$$\|r_n\| = |\langle t_{n+1}, z_n\rangle| \, \|r_{n-1}\|$$

and

$$\langle t_{n+1}, z_n\rangle = \langle t_{n+1}, At_n\rangle \langle t_n, A^{-1} z_n\rangle.$$

Proof. If $r_{n-1} = 0$, everything cancels in the previous expression (r_n, z_n and t_{n+1}). If A_n is singular then $\|r_n\| = \|r_{n-1}\|$ (see the discussion above) but also $|\langle t_{n+1}, z_n\rangle| = 1$, because these elements are parallel and have unit norm.

We then have to prove that, if A_n is regular (and $r_{n-1} \neq 0$), then

$$|\langle t_{n+1}, z_n\rangle|^2 = \frac{\|\widehat{r}_n\|^2}{\|r_{n-1}\|^2 + \|\widehat{r}_n\|^2}.$$

This result is a consequence of the following facts: (a) $\widehat{r}_n - r_{n-1} = A(u_{n-1} - y_n)$ belongs to $A\mathcal{W}^{(n)}$ and is orthogonal to $A\mathcal{W}^{(n-1)}$ and hence

$$\widehat{r}_n - r_{n-1} \in \mathbb{C}\langle z_n \rangle;$$

(b) $\widehat{r}_n = f - Ay_n \in \mathcal{W}^{(n+1)}$ and is orthogonal (by definition) to $\mathcal{W}^{(n)}$, so

$$\widehat{r}_n \in \mathbb{C}\langle t_{n+1} \rangle;$$

(c) t_{n+1} is orthogonal to $r_{n-1} \in \mathcal{W}^{(n)}$. Then

$$\begin{aligned} \|\widehat{r}_n\| &\overset{\text{(b)}}{=} |\langle t_{n+1}, \widehat{r}_n \rangle| \\ &\overset{\text{(c)}}{=} |\langle t_{n+1}, \widehat{r}_n - r_{n-1} \rangle| \\ &\overset{\text{(a)}}{=} |\langle t_{n+1}, z_n \rangle| \, \|\widehat{r}_n - r_{n-1}\| \\ &\overset{\text{(b,c)}}{=} |\langle t_{n+1}, z_n \rangle| \left(\|\widehat{r}_n\|^2 + \|r_{n-1}\|^2 \right)^{1/2}. \end{aligned}$$

For the last part, we notice that $A^{-1} z_n \in \mathcal{W}^{(n)}$ and that $\{t_1, \ldots, t_k\}$ is an orthonormal basis for $\mathcal{W}^{(k)}$, so

$$A^{-1} z_n - \langle A^{-1} z_n, t_n \rangle t_n \in \mathcal{W}^{(n-1)}$$

and

$$z_n - \langle A^{-1} z_n, t_n \rangle \, A t_n \in A\mathcal{W}^{(n-1)} \subset \mathcal{W}^{(n)}.$$

Hence

$$\langle t_{n+1}, z_n \rangle = \langle t_{n+1}, A t_n \rangle \langle t_n, A^{-1} z_n \rangle.$$

$\square$

5 Equations of the second kind

Proposition *If $A = \lambda I + K$ with $\lambda \neq 0$ and K compact, then:*

(a) $|\langle t_{n+1}, z_n \rangle| \to 0$.

(b) *A_n is invertible for n large enough.*

(c) $u_n \to u$.

Proof. By the previous results,

$$|\langle t_{n+1}, z_n \rangle| \leq \|A^{-1}\| \, |\langle t_{n+1}, A t_n \rangle|$$

but (see the recursion to compute t_{n+1}

$$|\langle t_{n+1}, A t_n \rangle| = \|(I - P_n) A t_n\| = \|(I - P_n) K t_n\| \leq \|K t_n\|.$$

Since t_n is an orthonormal sequence, then $K t_n \to 0$, which proves (a). This proves that $|\langle t_{n+1}, z_n \rangle| < 1$ for n large enough and then A_n has to be regular.

Finally $0 \le \|r_n\| \le \|r_{n-1}\|$ and

$$\frac{\|r_{n-1}\|}{\|r_n\|} = |\langle t_{n+1}, z_n\rangle| \to 0$$

which implies that $\|r_n\| \to 0$ (it has to converge and it cannot converge to a positive value). $\square$

Proposition *For all n*

$$\|r_n\| \le \|A^{-1}\|^n \sigma_1(K)\dots\sigma_n(K)\|f\|.$$

Proof. It is clear that

$$\begin{aligned}
\|r_n\| &= |\langle t_{n+1}, z_n\rangle|\,\|r_{n-1}\| \\
&= |\langle t_{n+1}, z_n\rangle \dots \langle t_2, z_1\rangle|\ \|f\| \\
&\le \|A^{-1}\|^n\ |\langle t_{n+1}, At_n\rangle, \dots, \langle t_2, At_1\rangle|\ \ \|f\|.
\end{aligned}$$

We now also that

$$\langle t_{j+1}, At_i\rangle = \langle t_{j+1}, Kt_i\rangle$$

and that

$$At_i \in A\mathcal{W}^{(i)} \subset \mathcal{W}^{(i+1)} \perp t_{j+1}, \qquad j \ge i$$

son the matrix $\langle t_{j+1}, At_i\rangle$ is lower triangular. The corresponding finite dimensional operator is $K_n := P_n K(P_{n+1} - P_1)$ and

$$|\langle t_{n+1}, At_n\rangle, \dots, \langle t_2, At_1\rangle| = |\det(\langle t_{j+1}, At_i\rangle)| = \sigma_1(K_n)\dots\sigma_n(K_n) \le \sigma_1(K)\dots\sigma_n(K).$$

6 Preconditioned Petrov–Galerkin methods

Problem. Let $V : H \to \widehat{H}$ be invertible and $K : H \to \widehat{H}$ be compact and assume that $V + K$ is invertible (which is equivalent to its being one–to–one). We consider two sequences (directed in a parameter $h \to 0$) of finite dimensional subspaces

$$H_h \subset H, \qquad \widehat{H}_h \subset \widehat{H}, \qquad \dim H_h = \dim \widehat{H}_h$$

and the discrete equations

$$\left|\begin{array}{l} u_h \in H_h, \\[2ex] \langle (V+K)u_h, r_h\rangle = \langle f, r_h\rangle, \qquad \forall r_h \in \widehat{H}_h. \end{array}\right.$$

We assume that the Petrov–Galerkin method given above is convergent, i.e. $u_h \to u = (V+K)^{-1}f$ (for all f). Let $Q_h : \widehat{H} \to \widehat{H}_h$ be the orthogonal projection onto $\widehat{H}_h$. Notice that

$$Q_h(V+K)u_h = Q_h f.$$

Let finally $V_h := Q_h V|_{H_h}$. We intend to solve with GMRES the *preconditioned* system

$$u_h + V_h^{-1}Q_h K u_h = Q_h f.$$

Proposition *There exists $\beta > 0$ independent of h and f such that*

$$\|r_n^h\| \le \beta^n \sigma_1(K) \dots \sigma_n(K) \|f\|.$$

Proof. From the convergence of the method (and the independence of this concept with respect to compact perturbations), it follows that there exists $\alpha > 0$ such that

$$\sup_{0 \neq u_h \in H_h} \frac{|\langle V u_h, r_h \rangle|}{\|u_h\|} \ge \alpha \|r_h\|, \qquad \forall r_h \in \widehat{H}_h,$$

$$\sup_{0 \neq u_h \in H_h} \frac{|\langle (V+K) u_h, r_h \rangle|}{\|u_h\|} \ge \alpha \|r_h\|, \qquad \forall r_h \in \widehat{H}_h$$

If $K_h := Q_h K|_{H_h} : H_h \to \widehat{H}_h$, then

$$\|V_h^{-1}\| \le 1/\alpha, \qquad \|(V_h + K_h)^{-1}\| \le 1/\alpha.$$

Let $R_h := V_h^{-1} Q_h K : H \to H$, which is compact. Then we are applying GMRES to the infinite dimensional system

$$(I + R_h) u_h = Q_h f.$$

(However, the method will stop after at most $N = \mathrm{dim} H_h$ iterations, since it is equivalent to a finite dimensional system). Then

$$\|r_n^h\| \le \|(I + R_h)^{-1}\|^n \sigma_1(R_h) \dots \sigma_n(R_h) \|Q_h f\|.$$

To end the proof, we simple note that: (a) $\|Q_h f\| \le \|f\|$; (b) for all j and h

$$\sigma_j(R_h) \le \|V_h^{-1}\| \, \|Q_h\| \, \sigma_j(K) \le (1/\alpha) \sigma_j(K);$$

(c) for all h

$$\|(I + R_h)^{-1}\| = \|(V_h + K_h)^{-1} V_h\| \le \|V\| / \alpha.$$

□

7 Appendix

7.1 Singular values of compact operators

Proposition (Rayleigh quotients) *Let V be a compact self–adjoint positive operator:*

$$V := \sum_{n=1}^{\infty} \lambda_n \langle \cdot \, , \phi_n \rangle \phi_n,$$

with $\lambda_n \ge \lambda_{n+1} > 0$ and ϕ_n orthonormal. Let

$$T_n := \mathbb{C} \langle \phi_1, \dots, \phi_n \rangle.$$

Then

$$\sup_{0 \neq \psi \in T_n^{\perp}} \frac{\langle V\psi, \psi \rangle}{\|\psi\|^2} = \lambda_{n+1} \le \sup_{0 \neq \psi \in X_n^{\perp}} \frac{\langle V\psi, \psi \rangle}{\|\psi\|^2}$$

if $\dim X_n \le n$.

Proof. The first equality is straightforward. For the second one, take $0 \neq \psi \in T_{n+1} \cap X_n^{\perp}$.
□

Singular value decomposition. $K : H \to \widehat{H}$ is compact if and only if

$$K = \sum_{j=1}^{\infty} \sigma_j \, \langle \,\cdot\, , \phi_j \rangle \psi_j$$

with $\sigma_j > 0$ non–increasing and $\{\phi_j\}$, $\{\psi_j\}$ orthonormal. We denote $\sigma_n(K)$ to the singular values of K. Notice that for all subspace X_n such that $\dim X_n \leq n$

$$\sigma_{n+1}(K) \leq \sup_{0 \neq \phi \in X_n^{\perp}} \frac{|(K\phi, K\phi)|}{\|\phi\|^2}$$

with equality attained with $T_n := \mathbb{C}\langle \phi_1, \ldots, \phi_n \rangle$.

Proposition *If K is compact and A is bounded*

$$\sigma_n(AK) \leq \|A\| \, \sigma_n(K), \qquad \forall n.$$

and

$$\sigma_n(KA) \leq \|A\| \, \sigma_n(K), \qquad \forall n.$$

Proof. For the first singular value we can bound

$$\sigma_1(AK)^2 = \sup_{0 \neq \phi \in H} \frac{|(AK\phi, AK\phi)|}{\|\phi\|^2} \leq \|A\|^2 \sup_{0 \neq \phi \in H} \frac{|(K\phi, K\phi)|}{\|\phi\|^2} = \|A\|^2 \sigma_1(K)^2.$$

For the remaining ones

$$\sigma_{n+1}(K)^2 = \sup_{0 \neq \phi \in T_n^{\perp}} \frac{|(K\phi, K\phi)|}{\|\phi\|^2} \geq \|A\|^{-2} \sup_{0 \neq \phi \in T_n^{\perp}} \frac{|(AK\phi, AK\phi)|}{\|\phi\|^2} \geq \|A\|^{-2} \sigma_{n+1}(AK)^2.$$

Finally

$$\sigma_n(KA) = \sigma_n((KA)^*) = \sigma_n(A^*K^*) \leq \|A^*\| \, \sigma_n(K^*) = \|A\| \, \sigma_n(K).$$

□

Proposition *If A is an isomorphism and K is compact, then*

$$(1/\|A^{-1}\|)\sigma_n(K) \leq \sigma_n(AK) \leq \|A\|\sigma_n(K).$$

Proof. It is a simple consequence of the previous result:

$$\sigma_n(K) = \sigma_n(A^{-1}AK) \leq \|A^{-1}\| \, \sigma_n(AK).$$

□

7.2 Generalized SVD

Riesz bases. A sequence $\{\psi_n\}$ is a Riesz basis for H if there exists an isomorphism $A : H \to H$ and a Hilbert basis $\{\phi_n\}$ such that

$$\psi_n = A\phi_n, \qquad \forall n.$$

Proposition *Let $\{\psi_n\}$ be a Riesz basis of H. Then:*

(a) *There exist $C_1, C_2 > 0$ such that*

$$C_1\|u\|^2 \leq \sum_{n=1}^{\infty} |\langle u, \psi_n\rangle|^2 \leq C_2\|u\|^2, \qquad \forall u \in H.$$

(b) *$\mathbb{C}\langle \psi_n \,|\, n \geq 1\rangle$ is dense in H.*

(c) *There exists another Riesz basis $\{\eta_n\}$ (conjugate Riesz basis) such that*

$$\langle \psi_m, \eta_n\rangle = \delta_{nm}$$

Moreover, for all $u \in H$,

$$u = \sum_{n=1}^{\infty} \langle u, \eta_n\rangle\, \psi_n.$$

Proof. Let $\psi_n = A\phi_n$ with A isomorphism and ϕ_n Hilbert basis. Then

$$\sum_n |\langle u, \psi_n\rangle|^2 = \sum_n |\langle A^* u, \phi_n\rangle|^2 = \|A^* u\|^2, \qquad \forall u \in H.$$

The first result holds then with $C_1 = 1/\|A^{-1}\|$ and $C_2 = \|A\|$. The second one is straightforward. The conjugate basis is defined by

$$\eta_n := (A^{-1})^* \phi_n$$

and

$$u = A\,A^{-1} u = A\left(\sum_n \langle A^{-1}u, \phi_n\rangle \phi_n\right) = \sum_n \langle u, (A^{-1})^*\phi_n\rangle A\phi_n,$$

which proves the results. $\square$

Proposition *$K : H \to \widehat{H}$ is compact if and only if there exist two Riesz bases $\{\phi_n\}$ and $\{\psi_n\}$ (of H and $\widehat{H}$ respectively) and a sequence of positive non–increasing values $\sigma_n \to 0$ such that*

$$K = \sum_{n=1}^{\infty} \sigma_n \langle\,\cdot\,, \phi_n\rangle \psi_n.$$

Moreover, there exist $\alpha_1, \alpha_2 > 0$ such that

$$\alpha_1\sigma_n \leq \sigma_n(K) \leq \alpha_2\sigma_n, \qquad \forall n.$$

Proof. Equivalence of compactness and the series form above is straightforward. There exist isomorphisms $A : H \to H$ and $B : \widehat{H} \to \widehat{H}$ such that $A^{-1}\phi_n =: \widetilde{\phi}_n$ and $B^{-1}\psi_n =: \widetilde{\psi}_n$ are Hilbert bases. Then $K = BK_0A$ where

$$K_0 = \sum_n \sigma_n \langle \,\cdot\,, \widetilde{\phi}_n \rangle \widetilde{\psi}_n, , \qquad \sigma_n = \sigma_n(K_0),$$

and

$$(1/\|A^{-1}\|\,\|B^{-1}\|)\sigma_n \leq \sigma_n(AK_0B) = \sigma_n(K) \leq \|A\|\,\|B\|\,\sigma_n.$$

□

7.3 Singular values of bounded operators

Proposition *Let*

$$K = \sum_n \sigma_n \langle \,\cdot\,, \phi_n \rangle\, \psi_n$$

be a compact operator (given by its SVD). Then

$$\sigma_{n+1}(K) = \|K - K_n\|, \qquad K_n := \sum_{j=1}^{n} \sigma_j \langle \,\cdot\,, \phi_j \rangle\, \psi_j$$

and

$$\sigma_{n+1}(K) \leq \|K - C_n\|$$

for all C_n linear such that $\dim \mathcal{R}(C_n) \leq n$.

Proof. If $\dim \mathcal{R}(C_n) \leq n$, there exists $0 \neq \phi \in \mathbb{C}\langle \phi_1, \ldots, \phi_{n+1} \rangle$ such that $C_n\phi = 0$. Then

$$\|K\phi - C_n\phi\|^2 = \|K\phi\|^2 = \sum_{j=1}^{n+1} \sigma_j |\langle \phi, \phi_j \rangle|^2 \leq \sigma_{n+1}^2 \|\phi\|^2$$

and therefore $\|K - C_n\| \geq \sigma_{n+1}$. □

Definition. Let A be bounded. We define for all $n \geq 0$

$$\sigma_{n+1}(A) := \inf\{\|A - A_n\| \,|\, \dim \mathcal{R}(A_n) \leq n\}.$$

Notice that $\sigma_1(A) = \|A\|$ and

$$0 \leq \sigma_n(A) \leq \sigma_{n-1}(A) \leq \|A\|, \qquad \forall n.$$

Moreover $\sigma_n(A) \to 0$ if and only if A is compact. Otherwise

$$\sigma_n(A) \to \sigma_\infty > 0.$$

Finally $\sigma_n(A) = \sigma_n(A^*)$.

Proposition *If A is an isomorphism, then*

$$1/\|A^{-1}\| \leq \sigma_n(A) \leq \|A\|, \qquad \forall n.$$

Therefore, if A is an isometric isomorphism, then $\sigma_n(A) = 1$ for all n.

Proof. If $\dim \mathcal{R}(A_n) < \infty$, we take

$$0 \neq v \in \mathcal{R}(A_n A^{-1})^{\perp}, \qquad u = A^{-1}v$$

and notice that

$$\begin{aligned} \|Au - A_n u\|^2 &= \|v - A_n A^{-1} v\|^2 \\ &= \|v\|^2 + \|A_n A^{-1} v\|^2 \\ &\geq \|v\|^2 = \|Au\|^2 \geq (1/\|A^{-1}\|^2)\|u\|^2. \end{aligned}$$

Therefore $\|A - A_n \geq 1/\|A^{-1}\|$ for all A_n with finite rank. ◻

Proposition *Let A and B be bounded operators. Then*

$$\sigma_n(AB) \leq \|A\|\sigma_n(B), \qquad \forall n.$$

Proof. If B_n has $n-$dimensional range, then

$$\|A\|\,\|B - B_n\| \geq \|A\,B - A\,B_n\| \geq \sigma_{n+1}(AB)$$

and we can take the infimum in the left–hand side. ◻

Proposition *If A is invertible*

$$(1/\|A\|^2)\sigma_n(A) \leq \sigma_n(A^{-1}) \leq \|A^{-1}\|^2\sigma_n(A), \qquad \forall n.$$

Proof. Let $\dim\mathcal{R}(A_n) \leq n$. Then $\dim\mathcal{R}(A^{-1}A_nA^{-1}) \leq n$ and

$$\sigma_{n+1}(A^{-1}) \leq \|A^{-1} - A^{-1}A_nA^{-1}\| \leq \|A^{-1}\|\,\|A - A_n\|\,\|A^{-1}\|$$

and taking the infimum we obtain

$$\sigma_{n+1}(A^{-1}) \leq \|A^{-1}\|^2\,\sigma_{n+1}(A).$$

◻

Remark. If K is compact

$$\sigma_n(I + K) \leq 1 + \sigma_n(K).$$

Proof. Taking K_n (the $n-$th section of K) we prove that

$$\sigma_{n+1}(I + K) \leq \|I + K - K_n\| \leq 1 + \|K - K_n\| = 1 + \sigma_{n+1}(K).$$

◻

7.4 Finite rank operators and matrices

Situation. Consider two finite dimensional spaces H_N and $\widehat{H}_M$ and respective orthonormal bases of them: $\{\widetilde{\phi}_i\}$ and $\{\widetilde{\psi}_i\}$. Consider the matrix

$$\mathbf{A} = (a_{ij}) = (\langle A\widetilde{\phi}_i, \widetilde{\psi}_j\rangle)$$

Proposition *For all n*

$$\sigma_n(A) = \sigma_n(\mathbf{A}).$$

Proof. Consider the SVD of $\mathbf{A} = \mathbf{Q}\,\Sigma\mathbf{P}^*$. It is simple to see that for all $u \in H_N$

$$\begin{aligned} Au &= \sum_j \langle Au, \widetilde{\psi}_j\rangle \widetilde{\psi}_j \\ &= \sum_{i,j} \langle A\widetilde{\phi}_i, \widetilde{\psi}_j\rangle \langle u, \widetilde{\phi}_i\rangle \widetilde{\psi}_j \\ &= \sum_{i,j} a_{ij} \langle u, \widetilde{\phi}_i\rangle \widetilde{\psi}_j \\ &= \sum_{i,j,k} \langle u, \widetilde{\phi}_i\rangle \overline{p_{ik}} \sigma_k q_{kj} \widetilde{\psi}_j \\ &= \sum_k \langle u, \phi_k\rangle \sigma_k \psi_k. \end{aligned}$$

where

$$\psi_k := \sum_j q_{kj}\widetilde{\psi}_j, \qquad \phi_k := \sum_i p_{ki}\widetilde{\phi}_i$$

are orthonormal bases. The last expression is the SVD for the operator A. □

Remark.. For every square matrix

$$|\det \mathbf{A}| = \sigma_1(\mathbf{A}) \dots \sigma_n(\mathbf{A}).$$

References

A great deal of what's in here is based on materials that appear in these two papers.

Moret, Igor A note on the superlinear convergence of GMRES. SIAM J. Numer. Anal. 34 (1997), no. 2, 513–516

Moret, Igor Discrete Krylov subspace methods for equations of the second kind. Int. J. Comput. Math. 69 (1998), no. 3-4, 351–369