Helsinki University of Technology Institute of Mathematics Research Reports
Teknillisen korkeakoulun matematiikan laitoksen tutkimusraporttisarja
Espoo 2000 A426

A REFINED POLAR DECOMPOSITION: $A = UPD$

Timo Eirola

TEKNILLINEN KORKEAKOULU
TEKNISKA HÖGSKOLAN
HELSINKI UNIVERSITY OF TECHNOLOGY
TECHNISCHE UNIVERSITÄT HELSINKI
UNIVERSITE DE TECHNOLOGIE D'HELSINKI

Timo Eirola: A refined polar decomposition: $A = UPD$; Helsinki University of Technology Institute of Mathematics Research Reports A426 (2000).

Abstract: *A refinement of the polar decomposition of a nonsingular matrix* A *is considered. Here* A *is written as a product of unitary* U, *Hermitian and positive definite* P *which has unit diagonal, and diagonal positive* D. *It is shown that such a decomposion exists and is unique. Rectangular and singular cases are also considered. Then a simple fixed point iteration using SVD is given to compute this decomposition. Also implementation of the Newton's method is discussed*
The refined polar decomposition can be used to parameterize the orbit of a matrix with distinct eigenvalues.

AMS subject classifications: 15A23, 65F35.

Keywords: Polar decomposition, orbits of matrices.

ISBN 951-22-4847-6
ISSN 0784-3143
Intitute of Mathematics, Espoo, 2000

Helsinki University of Technology
Department of Engineering Physics and Mathematics
Institute of Mathematics
P.O. Box 1100, 02015 HUT, Finland
email: *math@hut.fi*
downloadables: *http://www.math.hut.fi/*

author's email: *Timo.Eirola@hut.fi*

1 Introduction

The polar decomposition of a matrix $A \in \mathbb{C}^{m\times n}$, $m \geq n$, is given as

$$A = QS \ ,$$

where $Q \in \mathbb{C}^{m\times n}$ has orthonormal columns and $S \in \mathbb{C}^{n\times n}$ is Hermitian and positive semidefinite. S is unique and, in the case A has rank n, also Q is unique (see, e.g., [3],[5]). They are given by

$$S = (A^*A)^{1/2} \ , \qquad Q = A(A^*A)^{-1/2} \ , \tag{1.1}$$

where $^{1/2}$ denotes the positive semidefinite square root.

Here we want to further decompose $S = VPD$ so that this decomposition — with unitary V and nonnegative diagonal D — would have P Hermitian positive semidefinite and the diagonal elements of P should be ones. This leads to the decomposition

$$A = UPD \ , \tag{1.2}$$

where $U \in \mathbb{C}^{m\times n}$ has orthonormal columns, $P \in \mathbb{C}^{n\times n}$ is Hermitian and positive semidefinite with unit diagonal, and $D \in \mathbb{R}^{n\times n}$ is diagonal and nonnegative.

In section 2 it is shown that the decomposition (1.2) always exists. D is unique and if A does not have columns that are zero, then P is unique, too, and so is U if A has full rank. In the set of full rank matrices these factors are smooth functions of A.

Then in section 3 two methods are considered for computing (1.2). First a simple fixed point iteration of the D part is given and is shown to be locally convergent. Then two variants of the Newton iteration are discussed.

Finally two applications are given. The first is to parameterize the orbit (set of similar matrices) of a matrix with distinct eigenvalues, the second is almost optimal diagonal scaling to reduce the condition number.

2 Existence and uniqueness

In this section the main theorem concerning (1.2) is given. For this note first that the polar decomposition is real analytic as a function of full rank A. This is seen from (1.1) and

$$S = \tfrac{1}{2\pi i}\int_\gamma \sqrt{z}\,(z - A^*A)^{-1}\,dz \ , \tag{2.1}$$

where γ is a positively oriented simple closed curve in the right hand side of $\mathbb{C}$ enclosing the spectrum of A^*A, and $\sqrt{z}$ denotes the root with positive real part ([5], see also [1]). Formula (2.1) holds even in the case of singular A now γ enclosing only the nonzero eigenvalues. It follows that S is smooth in each of the sets of constant rank matrices.

Notation: $|v|$ denotes the 2–norm of $v \in \mathbb{C}^n$, $\mathcal{D} \subset \mathbb{C}^{n\times n}$ is the set of diagonal matrices, $\mathcal{D}_+ \subset \mathcal{D}_\mathbb{R} \subset \mathcal{D}$ contain the ones with, respectively, nonnegative and real entries, and $\mathrm{Diag}(M) \in \mathcal{D}$ is the diagonal of $M \in \mathbb{C}^{n\times n}$ and $\mathrm{diag}(M) \in \mathbb{C}^n$ is the corresponding vector. For $x \in \mathbb{R}^n$ denote $\mathbf{D}(x) = \begin{bmatrix} x_1 & & \\ & \ddots & \\ & & x_n \end{bmatrix} \in \mathcal{D}_\mathbb{R}$ and let $\mathbf{X} : \mathcal{D}_\mathbb{R} \to \mathbb{R}^n$ be its inverse mapping. $A \circ B$ will denote the Hadamard (elementwise) product.

Theorem 2.1 (Refined polar decomposition). *Assume* $A \in \mathbb{C}^{m\times n}$, $m \geq n$. *Then there exists a decomposition*

$$A = UPD\ , \tag{2.2}$$

where $U^*U = I$, P *is Hermitian, positive semidefinite,* $\mathrm{Diag}(P) = I$, *and* $D \in \mathcal{D}_+$.

D *is unique. If* A *does not have zero columns, then* P *is unique. If* $\mathrm{rank}(A) = n$, *then* U *is also unique and* P, D *are positive definite.*

Proof. Assume first that A does not have a zero column.

Existence. For given positive diagonal D let $U_D P_D = AD$ be a polar decomposition of AD. Then P_D is unique. Set $F(D) = \mathrm{Diag}(P_D)$. By (2.1) F is continuous. Then the existence of (2.2) amounts to solving $F(D^{-1}) = I$.

Let $a_1, \dots, a_n$ and $p_1, \dots, p_n$, respectively, be the columns of A and P_D. These are nonzero. Set

$$m = \min_j |a_j|\ , \quad M = \max_j |a_j|\ .$$

Then $p_{ii} \leq |p_i| = |a_i| d_i \leq M d_i$. Since P_D is Hermitian and positive semidefinite, for all i, j holds $p_{ii} p_{jj} \geq |p_{ij}|^2$. Thus $M d_i\, p_{jj} \geq |p_{ij}|^2$ and

$$M \left(\sum_i d_i\right) p_{jj} \geq \sum_i |p_{ij}|^2 = |a_j|^2 d_j^2 \geq m^2 d_j^2\ .$$

Hence for any $D = \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{bmatrix} \in \mathcal{D}_+$ we get

$$\frac{m^2}{M} \frac{d_j^2}{\sum_i d_i} \leq F_j(D) \leq M d_j\ . \tag{2.3}$$

Fix $\alpha \in (0, 1/2]$. For $\delta \in \mathbb{R}^n$ set

$$g(\delta) = \delta + \alpha \mathbf{X}(\log(F(e^{-\mathbf{D}(\delta)}))) \in \mathbb{R}^n\ . \tag{2.4}$$

Assume $\delta_j \in [\lambda, \Lambda]$ for all j. Then inequalities (2.3) imply

$$g_j(\delta) \leq \delta_j + \alpha \log(M e^{-\delta_j}) = (1-\alpha)\delta_j + \alpha \log(M) \leq (1-\alpha)\Lambda + \alpha \log(M)$$

and

$$\begin{aligned} g_j(\delta) &\geq \delta_j + \alpha \log(m^2 e^{-2\delta_j}) - \alpha \log\left(M \ \textstyle\sum_i e^{-\delta_i}\right) \\ &\geq (1-2\alpha)\delta_j + \alpha \log(m^2) - \alpha \log(Mn) + \alpha\lambda \\ &\geq (1-\alpha)\lambda + \alpha \log(\tfrac{m^2}{Mn}) \ . \end{aligned}$$

Choosing

$$\lambda = \log(\tfrac{m^2}{Mn}) \qquad \text{and} \qquad \Lambda = \log(M)$$

we get $g_j(\delta) \in [\lambda, \Lambda]$ for all j. Hence g maps the convex cube $[\lambda, \Lambda]^n \subset \mathbb{R}^n$ into itself. Further, g is continuous. Hence, by the Brower's fixed point theorem (see, e.g., [7]), there exists $\delta \in \mathcal{D}_{\mathbb{R}}$ such that $g(\delta) = \delta$ and $\widetilde{D} = e^{\mathbf{D}(\delta)}$ solves $F(\widetilde{D}^{-1}) = I$.

Uniqueness. Define $f \, : \, \mathbb{R}^n \to \mathbb{R}^n$ as

$$f(d) = \mathbf{X}(F(e^{\mathbf{D}(d)})) \ . \tag{2.5}$$

By lemma 2.2 below, the derivative $f'(d) \in \mathbb{R}^{n \times n}$ is positive definite for all d. Hence, if $f(\widehat{d}) = f(d)$ then

$$0 = (\widehat{d} - d)^T (f(\widehat{d}) - f(d)) = (\widehat{d} - d)^T \int_0^1 f'(d + t(\widehat{d} - d)) \, dt \ (\widehat{d} - d)$$

so that $\widehat{d} - d = 0$, since the integral is positive definite. Thus f, and consequently also F, is an injection. This shows uniqueness of D. Then uniqueness of P and U follow from the corresponding properties of the polar decompostion.

If A has $q \geq 1$ columns that are zero then, necessarily, we set the corresponding elements of D to zero. Let $\widehat{A} \in \mathbb{C}^{m \times (n-q)}$ consist of the nonzero columns of A. Take a refined decomposition $\widehat{A} = \widehat{U}\widehat{P}\widehat{D}$ and put elements of these in the corresponding places of $U, P,$ and D. Fill the rest of P (except diagonal elements) with zeros, and U with orthonormal columns, orthogonal to $\widehat{U}$. $\square$

Remark 2.1. If A has just one zero column then P is still unique. This is because in P the norms of the columns corresponding to nonzero columns of A are the same as those of $\widehat{P}$. Hence, except the diagonal element, the row of P corresponding to the zero column of A has to be zero.

The following lemma was needed above and will be used again when considering the computation of the refined polar decomposition with the Newton's method.

Lemma 2.2. *Assume A has no zero columns. Let $UP = Ae^{\mathbf{D}(d)}$ be a polar decomposition and Let $P = V\mathbf{D}(\pi)V^*$ be an eigendecomposition of P. Then the derivative of f of (2.5) is given by*

$$f'(d)\delta = \operatorname{diag}(V(\Pi \circ (V^* \mathbf{D}(\delta) V))V^*) \ , \tag{2.6}$$

where $\Pi_{ij} = \frac{\pi_i^2+\pi_j^2}{\pi_i+\pi_j}$, $\Pi_{ij} = 0$ *if* $\pi_i + \pi_j = 0$. *Further,* $f'(d)$ *is positive definite.*

Proof. We have $P = (e^{\mathbf{D}(d)}A^*Ae^{\mathbf{D}(d)})^{\frac{1}{2}}$ and this is differentiable with respect to d. This is true even in the case of singular A since $e^{\mathbf{D}(d)}A^*Ae^{\mathbf{D}(d)}$ has constant rank and formula (2.1) can be applied.

For small $\delta \in \mathbb{R}^n$ we need a Hermitian Δ such that

$$(P+\Delta)^2 = e^{\mathbf{D}(d+\delta)}A^*Ae^{\mathbf{D}(d+\delta)} + O(\delta^2+\Delta^2) \ ,$$

i.e.,

$$\begin{aligned} P\Delta + \Delta P &= \mathbf{D}(\delta)e^{\mathbf{D}(d)}A^*Ae^{\mathbf{D}(d)} + e^{\mathbf{D}(d)}A^*Ae^{\mathbf{D}(d)}\mathbf{D}(\delta) \\ &= \mathbf{D}(\delta)P^2 + P^2\mathbf{D}(\delta) \ . \end{aligned}$$

Using $P = V\mathbf{D}(\pi)V^*$ we get

$$\mathbf{D}(\pi)\ V^*\Delta V + V^*\Delta V\ \mathbf{D}(\pi) = V^*\mathbf{D}(\delta)V\ \mathbf{D}(\pi)^2 + \mathbf{D}(\pi)^2\ V^*\mathbf{D}(\delta)V \ ,$$

i.e.,

$$(\pi_i+\pi_j)(V^*\Delta V)_{ij} = (\pi_i^2+\pi_j^2)(V^*\mathbf{D}(\delta)V)_{ij} \ .$$

If $\pi_i + \pi_j = 0$ we take $(V^*\Delta V)_{ij} = 0$. Hence $V^*\Delta V = \Pi \circ (V^*\mathbf{D}(\delta)V)$ and (2.6) follows.

Let $u = (1,\ldots,1)$. Positive definiteness is shown by

$$\begin{aligned} \delta^T f'(d)\delta &= \mathrm{tr}(\mathbf{D}(\delta)V(\Pi\circ(V^*\mathbf{D}(\delta)V))V^*) \\ &= \mathrm{tr}(V^*\mathbf{D}(\delta)V(\Pi\circ(V^*\mathbf{D}(\delta)V))) \\ &= \textstyle\sum_{i,j} \frac{\pi_i^2+\pi_j^2}{\pi_i+\pi_j}\ |(V^*\mathbf{D}(\delta)V)_{ij}|^2 \\ &\geq \textstyle\sum_{i,j} \frac{\pi_i+\pi_j}{2}\ |(V^*\mathbf{D}(\delta)V)_{ij}|^2 \\ &= \tfrac{1}{2}\,\mathrm{tr}(V^*\mathbf{D}(\delta)V((\pi u^T + u\pi^T)\circ(V^*\mathbf{D}(\delta)V))) \\ &= \tfrac{1}{2}\,\mathrm{tr}(V^*\mathbf{D}(\delta)V\mathbf{D}(\pi)V^*\mathbf{D}(\delta)V) + \tfrac{1}{2}\,\mathrm{tr}(V^*\mathbf{D}(\delta)VV^*\mathbf{D}(\delta)V\mathbf{D}(\pi)) \\ &= \mathrm{tr}(\mathbf{D}(\delta)P\mathbf{D}(\delta)) = \textstyle\sum_j p_{jj}\delta_j^2 \ . \end{aligned}$$

□

Remark 2.2. In the proof of the theorem the g–function was defined by (2.4) for $\alpha \in (0,1/2]$. Values $\alpha \in (1/2,1)$ also work. Then the lower bound becomes

$$g_j(\delta) \geq (1-2\alpha)\Lambda + \alpha\lambda + \alpha\log(\tfrac{m^2}{Mn})$$

and the choice

$$\lambda = \tfrac{1}{1-\alpha}\ ((1-2\alpha)\Lambda + \alpha\log(\tfrac{m^2}{Mn}))$$

works. In the numerical computations we will mostly use $\alpha \approx 2/3$.

Remark 2.3. The Π-matrix above is quite interesting: it is positive and has only one positive eigenvalue. It seems to be negative semidefinite in the subspace $\left\{v \,\middle|\, \sum_j v_j = 0\right\}$. This would further imply that $f'(d)$ is an M-matrix (see (3.1) below).

Remark 2.4. Since $\frac{\pi_i^2+\pi_j^2}{\pi_i+\pi_j} \leq \max(\pi_i, \pi_j)$ we get $\|f'(d)\| \leq \max_j \pi_j = \|P\|$. In practice, however, we usually observe $\|f'(d)\| \approx 2$.

Finally, for good matrices the decompostion is smooth (real analytic):

Proposition 2.3. *In the set of full rank matrices $A \in \mathbb{C}^{m\times n}$, $m \geq n$ the factors U, P and D are real analytic functions of the (real and imaginary parts of the) elements of A.*

Proof. By the implicit function theorem D solving $F(D^{-1}) = I$ depends smoothly on A (F' is invertible). Then $P = (D^{-1}A^*AD^{-1})^{\frac{1}{2}}$ and $U = AP^{-1}D^{-1}$ are also smooth. □

3 Numerical computation

Here we consider the two obvious approaches to compute the refined polar decomposition. First we consider fixed point iterations of the g-function (2.4). Then we will apply (2.6) in a Newton scheme and consider also more economic approximate Newton steps.

3.1 Fixed point iteration

A simple numerical method is obtained by iteration of the map g of (2.4) with $\alpha \in (0, 1)$. We use the (economy version) singular value decomposition to compute values of f. From the singular value decomposition

$$Q\,\mathbf{D}(\pi)\,V^* = A\,e^{-\mathbf{D}(d)}$$

we get $P = V\,\mathbf{D}(\pi)\,V^*$ already eigendecomposed (see Lemma 2.2) and

$$f(-d) = \mathrm{diag}(P) = (V \circ \bar{V})\pi\ .$$

For `matlab` ([6]) and `octave` ([2]) we can write

```
function [U,P,D]=UPD_F(A)

% This function computes the refined polar decomposition
% of A using the fixed point iteration.

sz=size(A); n=sz(2); alpha=2/3;
expd=ones(n,1); d=zeros(n,1);
err=1; k=0; tol=10^(-13);

while err > tol,
   [Q,E,V]=svd(A*diag(expd),0);
   f=log((V.*conj(V))*diag(E));
   d=d+alpha*f; expd=exp(-d);
   err=norm(f); k=k+1; end

U=Q*V';  P=V*E*V';  D=diag(1./expd);
```

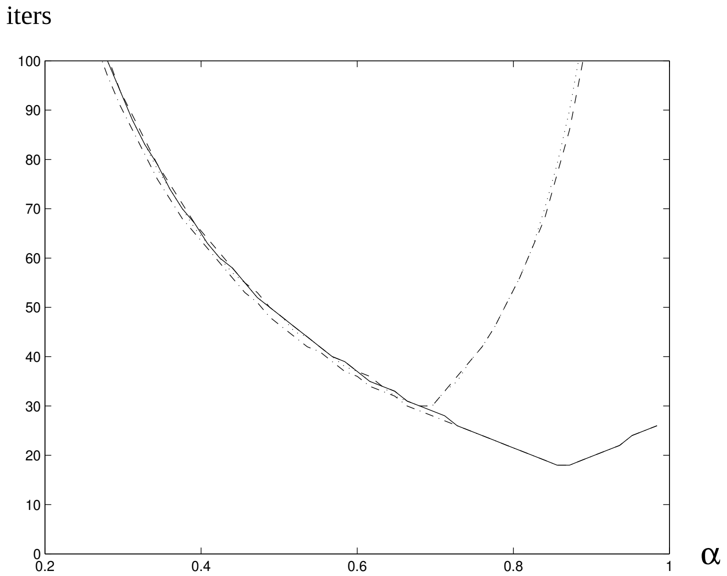
For small α local convergence of this iteration is guaranteed:

Proposition 3.1. *Given A, without zero columns, then the iteration $d^{k+1} = g(d^k)$ converges from d_0 close to $d = g(d)$ provided $\alpha \in (0, \frac{2}{\|P\|})$, where P is the Hermitian part of $A = UPD$.*

Proof. This follows directly from $g'(d) = I - \alpha f'(-d)$ and the positive definiteness of f'. Note that by remark 2.4 $\|f'(-d)\| \leq \|P\|$. $\square$

In the following figure α varies from 0.2 to 1 and the iteration counts for different matrices are plotted. The matrices are 50×50 random matrix (—) 20×20 Hilbert matrix $(\cdots)$, 50×50 random matrix of rank 25 (- - -), and a random 50×25 matrix $(-\cdot-\cdot-)$.

A thumb rule is that for S closer to a diagonal matrix α closer to one gives fastest convergence.



3.2 Newton iteration

We want to solve $f(d) = u$, where $u = (1, \ldots, 1)$. To obtain the polar decomposition for computing $f(d)$ we use again the singular value decomposition $Q\,\mathbf{D}(\pi)\,V^* = A\,e^{\mathbf{D}(d)}$ and $f(d) = (V \circ \bar{V})\pi$.

Using (2.6) we get the entries of f' :

$$\begin{aligned} f'(d)_{ij} &= e_i^T\, f'(d)\, e_j \\ &= \operatorname{tr}\left(\mathbf{D}(e_i)V\left(\Pi \circ (V^*\mathbf{D}(e_j)V)\right)V^*\right) \\ &= \operatorname{tr}\left(V^*\mathbf{D}(e_i)V\left(\Pi \circ (V^*\mathbf{D}(e_j)V)\right)\right) \\ &= \operatorname{tr}\left(v_i v_i^*\left(\Pi \circ (v_j v_j^*)\right)\right) \\ &= (v_i \circ \bar{v}_j)^*\Pi\,(v_i \circ \bar{v}_j)\ , \end{aligned} \tag{3.1}$$

where v_i's are the columns of V^*. The proof of Lemma 2.2 shows that

$$\operatorname{cond}(f'(d)) \leq \frac{\max_j \pi_j}{\min_i p_{ii}}\ .$$

At the solution this upper bound is $\leq n$. Hence good conditioning of the Jacobian and local convergence is guaranteed.

Due to (3.1), the complexity (flop count) is $O(n^4)$ per iteration step. On the other hand computing f' this way parallelizes easily.

In the experiments the initial guess $d = 0$ seems to work in most cases[1] but for safety we take first one step of the fixed point iteration of g with $\alpha = 2/3$ to obtain $d_0 = -g(0)$.

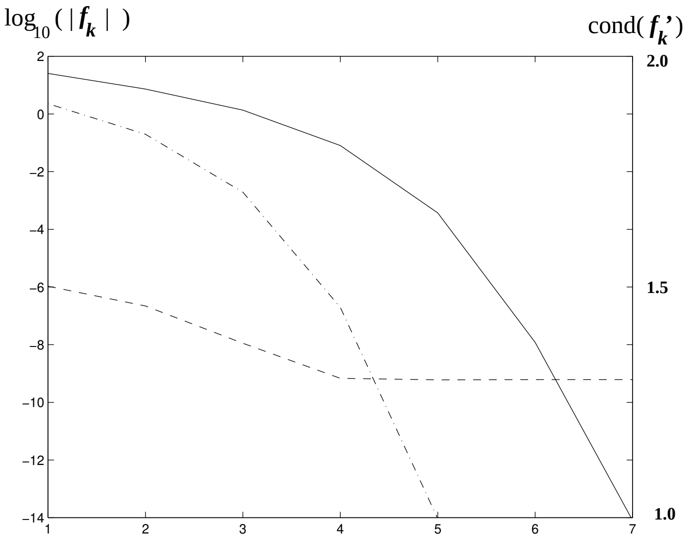
With these remarks the Newton's method for computing the refined polar decomposition can be written as follows

```
function [U,P,D]=UPD_N(A)

% This function computes the refined polar decomposition
% of A using the Newton's method.

sz=size(A); n=sz(2);
[U,E,V]=svd(A,0);
d=-2/3*log((V.*conj(V))*diag(E));
expd=exp(d); w=ones(n,1); df=zeros(n,n);
err=1; k=0; tol=10^(-13);
[U,E,V]=svd(A*diag(expd),0);
Vc=conj(V); p=diag(E);
f=(V.*Vc)*p-w;
```

[1] In the short series of test problems tried so far just a few cases required a better initial guess.

```
while err > tol,
   Pi=(p.^2*w'+w*p'.^2)./(p*w'+w*p');
   for i=1:n , for j=i:n ,                              % these
      v=(V(i,:).*Vc(j,:))';                             % take
      df(i,j)=v'*Pi*v;                                  % O(n^4)
      df(j,i)=df(i,j); end, end,                        % flops
   d=d-df\f; expd=exp(d);
   [Q,E,V]=svd(A*diag(expd),0);
   Vc=conj(V);   p=diag(E);
   f=real(V.*Vc)*p-w;
   err=norm(f); k=k+1; end

U=Q*V';  P=V*E*V';  D=diag(1./expd);
```

In the following figure a typical convergence graph is drawn. The solid line is 10–base logarithm of the norm of f when started from the trivial guess $d = 0$ and (–·–·–) corresponds to the better starting value. The dashed line plots cond(f'), Here A is a random 100×100 matrix (**randn(100)**).



3.3 Approximate Newton

In the Newton's method above the computation of f' is the most flops consuming part $O(n^4)$. In each Newton step we solve $f'(d)\delta = u - f(d)$, i.e. (see (2.6)),

$$\mathrm{diag}(V(\Pi \circ (V^*\mathbf{D}(\delta)V))V^*) = u - f(d) \ . \tag{3.2}$$

Let us eigendecompose $\Pi = W\Lambda W^T$. Π has one positive and many small negative eigenvalues. We take an approximation:

$$\Pi \approx \widetilde{\Pi} = \sum_{|\lambda_j|>\varepsilon} \lambda_j \ w_j w_j^T \ ,$$

where w_j 's are the columns of W. Using $\widetilde{\Pi}$ in (3.2) we get simplification:

$$
\begin{aligned}
&\operatorname{diag}\Big(V\big(\widetilde{\Pi}\circ(V^*\mathbf{D}(\delta)V)\big)V^*\Big)\\
&=\sum_{|\lambda_j|>\varepsilon}\lambda_j\ \operatorname{diag}\Big(V\big((w_jw_j^T\circ(V^*\mathbf{D}(\delta)V)\big)V^*\Big)\\
&=\sum_{|\lambda_j|>\varepsilon}\lambda_j\ \operatorname{diag}\Big((V\mathbf{D}(w_j)V^*)\ \mathbf{D}(\delta)\ (V\mathbf{D}(w_j)V^*)\Big)\\
&=\sum_{|\lambda_j|>\varepsilon}\lambda_j\ (G_j\circ\bar{G}_j)\ \delta\ ,
\end{aligned}
$$

where $G_j = V\mathbf{D}(w_j)V^*$. Hence:

$$
f'(d)\approx\sum_{|\lambda_j|>\varepsilon}\lambda_j\ (G_j\circ\bar{G}_j)\ .
$$

For $\varepsilon = 0.001$ we typically get 4-6 terms in the sum while the iteration count stays practically the same as for the genuine Newton's method. This takes the Newton step back to $O(n^3)$ as can be seen from the tests of the next section.

The `matlab/octave` code for this approximate Newton is obtained by replacing the four lines ("`these take O(n^4) flops`") by lines

```
[W,lambda]=eig(Pi); df=zeros(n,n);
for j=1:n ,
   if abs(lambda(j,j)) > rtol ,
      G=V*diag(W(:,j))*V';
      df=df+lambda(j,j)*(G.*conj(G)); end,end
```

Remark 3.1. Here we just use the `eig` routine to get the eigendecomposition of Π. Since we want only a couple of largest (in modulus) eigenvalues and the corresponding eigenvectors, the Lanczos iteration should give extra savings.

3.4 Comparison

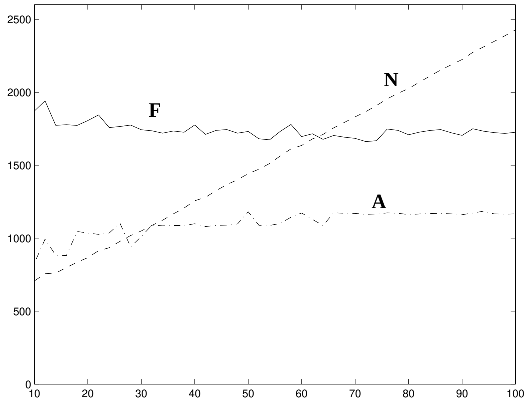
Below we have plotted the flop counts divided by n^3 of the three methods: F: the fixed point iteration with $\alpha = 2/3$ (—), N: the genuine Newton's method (- - -), and A: the approximate Newton's method (–·–·–).

We took four series of test problems. In each of these the (column) dimension (horizontal axis) grows from 10 to 100.

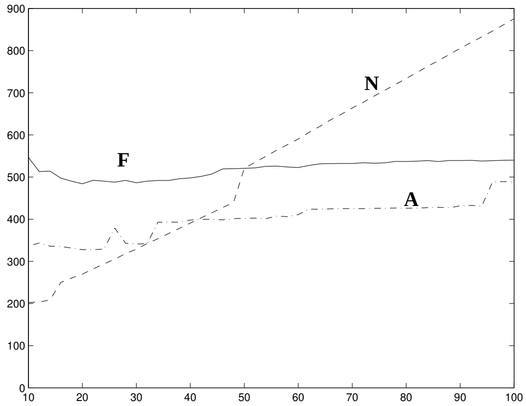
1. Complex random matrices (`A=randn(n)+sqrt(-1)*randn(n)`)
2. Hilbert matrices (`A=hilb(n)`)

3. Singular real matrices with $\mathrm{rank} = \frac{n}{2}$ (`A=randn(n,n/2)*randn(n/2,n)`)

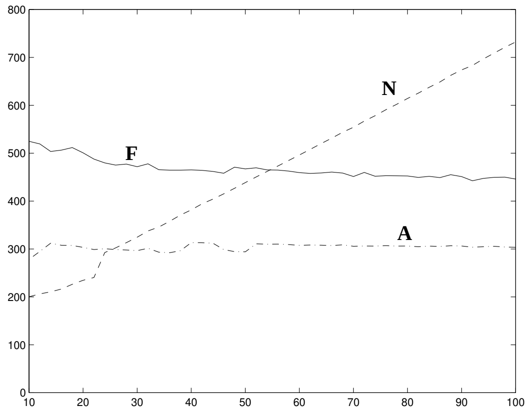
4. Random $2n \times n$ full rank matrices (`A=randn(2*n,n)`)



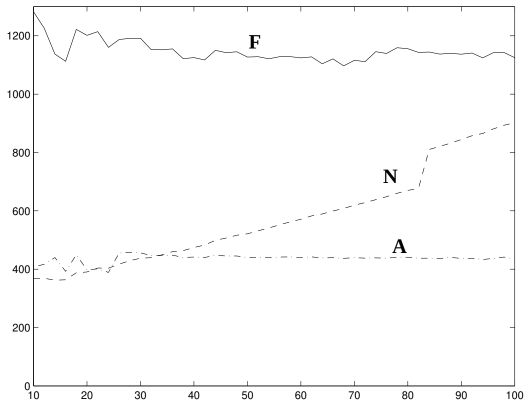
Random matrices



Hilbert matrices



Singular matrices: $\mathrm{rank}(A) = \frac{n}{2}$



Random $2n \times n$ full rank matrices

Methods F and A seem to have $O(n^3)$ complexity, while N is clearly $O(n^4)$.

Remark 3.2. The methods above are just simple first approaches. It will be interesting to study how the iterations for the polar decomposition (see e.g. [4]) can be adapted to this case.

4 Applications

4.1 Parameterizing the orbit of a diagonable matrix

The *orbit* of a matrix is the set of matrices similar to it.

Complex case. Let the eigenvalues of $A \in \mathbb{C}^{n\times n}$ be distinct. Then A is diagonable: $A = T\Lambda T^{-1}$, $\Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \in \mathcal{D}$, $\lambda_i \neq \lambda_j$ for $i \neq j$. The orbit of A is given by

$$\mathcal{S}(A) = \left\{ X\Lambda X^{-1} \,\middle|\, X \in \mathbb{C}^{n\times n}\,,\ \det(X) \neq 0 \right\}\ .$$

If X and Y are nonsingular such that $X\Lambda X^{-1} = Y\Lambda Y^{-1}$ then $Y^{-1}X\Lambda = \Lambda Y^{-1}X$ and for $i \neq j$ holds $(\lambda_i - \lambda_j)(Y^{-1}X)_{ij} = 0$. Hence $\widehat{D} = Y^{-1}X$ is diagonal. Write any diagonal nonsingular matrix as $\widehat{D} = ED$, where $|E_{ii}| = 1$ and $D_{ii} > 0$ for all i.

Let $X = UP$, i.e., U is unitary and $P \in \mathcal{P}_I$ – the set of Hermitian positive definite matrices with unit diagonal. Then all matrices that transform Λ to the same matrix as X does are of the form

$$Y = UPED\ .$$

Since $E^*PE \in \mathcal{P}_I$, too, and $\widehat{X} = UEE^*PE$ gives also $\widehat{X}\Lambda\widehat{X}^{-1} = X\Lambda X^{-1}$ we still have to choose E. We do this by requiring that the first nonzero entry in each column of U is real and positive. Let $\mathcal{U}$ denote the set of such unitary matrices. Then

$$\mathcal{S}(A) = \left\{ UP\Lambda P^{-1}U^* \,\middle|\, U \in \mathcal{U}\,,\ P \in \mathcal{P}_I \right\}$$

and for each $B \in \mathcal{S}(A)$ the factors U and P are uniquely defined. To separate the *unitary orbit* $\mathcal{S}_{\mathcal{U}}(A)$ and the transversal part $\mathcal{S}_{\mathcal{P}}(A)$ take the refined polar decomposition $T = U_0P_0D$ with $U_0 \in \mathcal{U}$. Then

$$\begin{aligned}
\mathcal{S}_{\mathcal{U}}(A) &= \left\{ UP_0\Lambda P_0^{-1}U^* \,\middle|\, U^*U = I \right\} \\
\mathcal{S}_{\mathcal{P}}(A) &= \left\{ U_0P\Lambda P^{-1}U_0^* \,\middle|\, P \in \mathcal{P}_I \right\}\ .
\end{aligned}$$

Real case. For $A \in \mathbb{R}^{n\times n}$ with distinct eigenvalues one might want to consider only the real orbit. If the eigenvalues are real, then we can proceed exactly as in the complex case, now only restricting to real matrices. This way we obtain unique coordinates for any real matrix in the orbit.

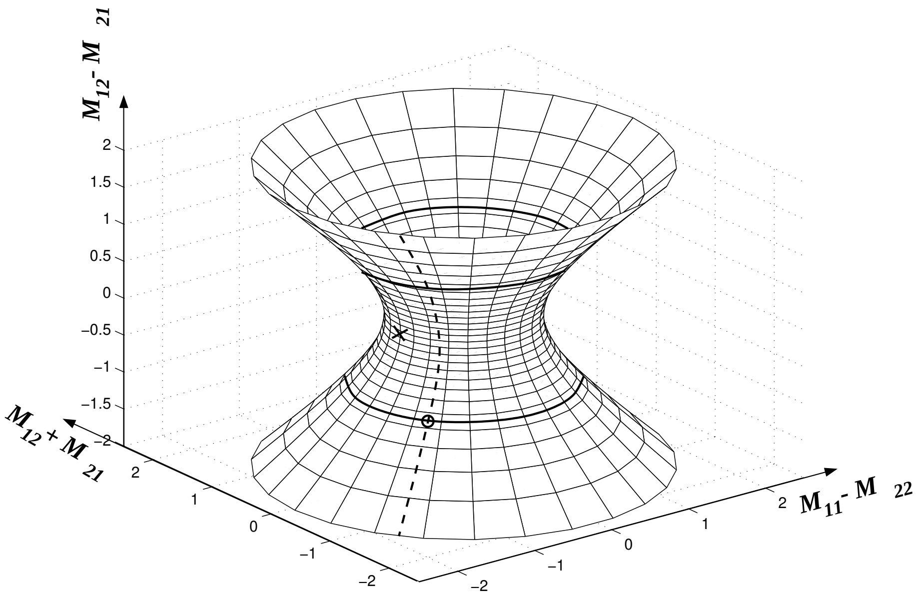
In the figure below the orbit of $A = \left[\begin{smallmatrix} 1.0 & -0.8 \\ 0.1 & 2.0 \end{smallmatrix}\right]$ is drawn. The displayed coordinates of $M \in \mathbb{R}^{2\times 2}$ are $M_{11} - M_{22}$, $M_{12} + M_{21}$ and $M_{12} - M_{21}$. The fourth coordinate $\operatorname{trace}(M)$ is not shown, since it is constant on orbits.

A is drawn as a small circle and $\Lambda = \left[\begin{smallmatrix} 1.0877 & 0 \\ 0 & 1.9123 \end{smallmatrix}\right]$ as a cross. The two darker circles on the surface form the *orthogonal orbit* of A :

$$\mathcal{S}_{\mathcal{O}}(A) = \left\{ UAU^T \,\middle|\, U \in \mathbb{R}^{n\times n}\,,\ U^TU = I \right\}\ .$$

The two parts correspond to orthogonal matrices with determinant ± 1, respectively. The dashed curve on the surface is the transversal part

$$\mathcal{S}_{\mathcal{P}}(A) = \left\{ U_0P\Lambda P^{-1}U_0^T \,\middle|\, P \in \mathcal{P}_I \cap \mathbb{R}^{n\times n} \right\}\ .$$



If $A \in \mathbb{R}^{n\times n}$ has distinct eigenvalues, but some of them are complex, then it admits a real similarity transformation $A = T\Lambda T^{-1}$ to real block diagonal Λ, where the blocks are either real numbers or 2×2 blocks of the form $\left[\begin{smallmatrix} \alpha & \beta \\ -\beta & \alpha \end{smallmatrix}\right]$.

Let $\mathcal{D}_\Lambda$ denote the set of block diagonal matrices with the same block structure as Λ has. Now the diagonal matrices do not commute with Λ but those in $\mathcal{D}_\Lambda$ do. Hence we want to consider refined polar decompositions

$$X = UP\widetilde{D}$$

with orthogonal U, symmetric positive definite P having unit diagonal, and $\widetilde{D} \in \mathcal{D}_\Lambda$ having nonnegative diagonal. Existence and uniqueness results can be obtained using similar techniques as in the proof of theorem 2.1. The idea is to write $\widetilde{D} = CD$, where $C, D \in \mathcal{D}_\Lambda$ and C is orthogonal and D is diagonal. This is to first transform a symmetric positive definite S to $\widetilde{S} = C^T SC$, so that the diagonal pairs of $\widetilde{S}$ corresponding to the 2×2 blocks of Λ match: $\widetilde{S}_{j,j} = \widetilde{S}_{j+1,j+1}$. Then combine this with diagonal scaling. This combination can then be used[2] in F.

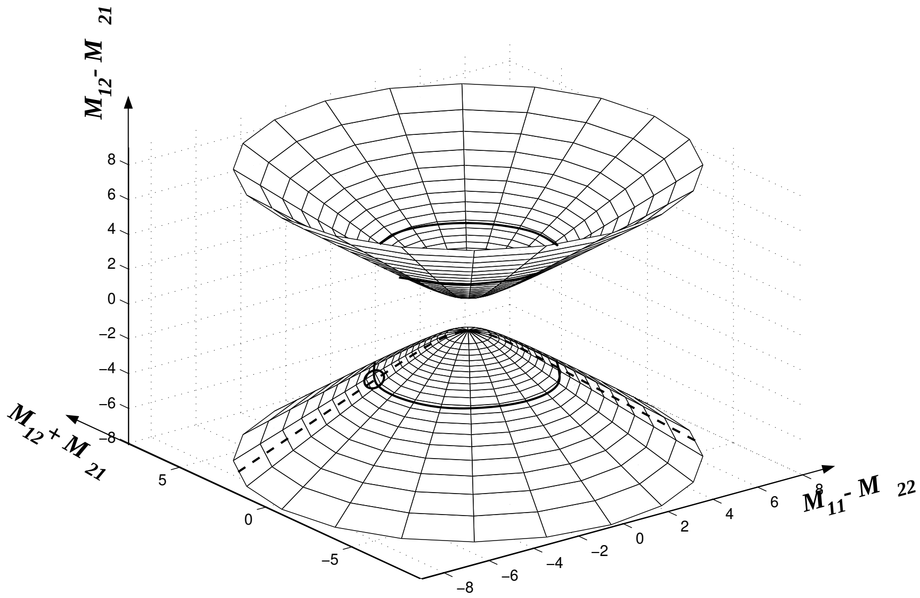
An algorithm for computing this is obtained by modifying the fixed point iteration. In the following code vector **z** contains the starting indices of the 2×2 blocks, i.e., it defines $\mathcal{D}_\Lambda$.

```
function [U,P,D]=C_UPD(A,z)

% This function computes the refined polar decomposition of A
% A=UPD with D "real C-diagonal" determined by z.
% Here the fixed point iteration is used.
```

[2]The details are not written here, since more general cases are under work.

```
sz=size(A); n=sz(2); alpha=2/3;
expd=ones(n,1); d=zeros(n,1); C=eye(n);
err=1; k=0; tol=10^(-13);

while err > tol,
   [U,E,V]=svd(A*diag(expd),0);
   for j=z , jj=j:j+1 ;
      W=V(jj,:)*E*V(jj,:)';
      fi=atan((W(1,1)-W(2,2))/(2*W(1,2)))/2;
      c=cos(fi); s=sin(fi); C(jj,jj)=[c,-s;s,c]; end
   V=C*V;
   f=log((V.*conj(V))*diag(E));
   d=d+alpha*f; expd=exp(-d);
   err=norm(f); k=k+1; end

U=U*V';  P=V*E*V';  D=C*diag(1./expd);
```

In the following figure the orbit of $A = \left[\begin{smallmatrix} -0.7 & -1.0 \\ 2.5 & 2.3 \end{smallmatrix}\right]$ is drawn. A is shown as a small circle and $\Lambda = \left[\begin{smallmatrix} 0.8 & -0.5 \\ 0.5 & 0.8 \end{smallmatrix}\right]$ is on the top of the lower part. The two darker circles on the surface form again the orthogonal orbit and the dashed curve is the transversal part.



Remark 4.1. The orthogonal orbits and transversal parts seem to intersect orthogonally (w.r.t. $\langle A, B\rangle = \operatorname{tr}(AB^*)$). This is true for 2×2 matrices, but not generally.

4.2 Diagonal scaling

Another way to use the refined polar decomposition is in diagonal scaling of a matrix to reduce its condition number. The following result shows that the diagonal part of the refined polar decomposition gives an almost optimal (right) scaling.

Proposition 4.1. *Let $A = UPD$ be the refined polar decomposition. Then for the 2-norm condition number*

$$\kappa(AD^{-1}) \leq n \inf_{E \in \mathcal{D}_+} \kappa(AE) \ .$$

Proof. A result of van der Sluis ([8]) says that if P is Hermitian positive definite and has constant diagonal, then

$$\kappa(P) \leq n \inf_{E \in \mathcal{D}_+} \kappa(EPE) \ .$$

Hence, denoting by $\lambda_{\min/\max}$, $\sigma_{\min/\max}$ the maximal and minimal eigenvalues and singular values:

$$\begin{aligned}
\kappa(AD^{-1}) = \kappa(UP) = \kappa(P) &\leq n \inf_{E \in \mathcal{D}_+} \kappa(EPE) \\
&= n \inf_{E \in \mathcal{D}_+} \frac{\lambda_{\max}(EPE)}{\lambda_{\min}(EPE)} = n \inf_{E \in \mathcal{D}_+} \frac{\lambda_{\max}(PE^2)}{\lambda_{\min}(PE^2)} \\
&\leq n \inf_{E \in \mathcal{D}_+} \frac{\sigma_{\max}(PE^2)}{\sigma_{\min}(PE^2)} = n \inf_{E \in \mathcal{D}_+} \kappa(PE) = n \inf_{E \in \mathcal{D}_+} \kappa(AE) \ .
\end{aligned}$$

□

Naturally, since the computation of the refined polar decomposition is so expensive, this result is at most a theoretical curiosity.

References

[1] L. Dieci and T. Eirola, *On smooth decompositions of matrices*, SIAM J. Matrix Anal. and Appl **20** (1999), 800–819.

[2] John W. Eaton, *Octave*, `www.che.wisc.edu/octave/`.

[3] G. Golub and C. Van Loan, *Matrix computations*, John Hopkins Univerity Press, Baltimore, 1989.

[4] N. Higham and R. Schreiber, *Fast polar decomposition of an arbitrary matrix*, SIAM J. Sci. Statist. Comput. **11** (1990), 648–655.

[5] R.A. Horn and C.R. Johnson, *Topics in matrix analysis*, Cambridge University Press, New York, 1991.

[6] Mathworks, *Matlab*, `www.mathworks.com/products/matlab/`.

[7] D. R. Smart, *Fixed point theorems*, Cambridge University Press, Cambridge, 1974.

[8] A. van der Sluis, *Condition numbers and equilibration of matrices*, Numer. Math. **14** (1969), 14–23.

(continued from the back cover)

A422 Marko Huhtanen
A matrix nearness problem related to iterative methods, Oct 1999

A421 Ph. Clément, G. Gripenberg, V. Högnäs, and S-O. Londen
Some remarks on the method of sums, Oct 1999

A420 Ville Havu and Juhani Pitkäranta
An Analysis of Finite Element Locking in a Parameter Dependent Model Problem, Aug 1999

A419 G. Gripenberg, Ph. Clément, and S-O. Londen
Smoothness in fractional evolution equations and conservation laws, Aug 1999

A418 Marko Huhtanen
Pole assignment problems for error bounds for GMRES, Aug 1999

A417 Kirsi Peltonen
Examples of uniformly quasiregular mappings, Jul 1999

A416 Saara Hyvönen
On the Iterative Solution of Nonnormal Problems, Jul 1999

A415 Ville Turunen
A smooth operator valued symbol analysis, Aug 1999

A414 Marko Huhtanen
A stratification of the set of normal matrices, May 1999

A413 Ph. Clément, G. Gripenberg and S-O Londen
Regularity properties of solutions of functional evolution equations, May 1999

A412 Marko Huhtanen
Ideal GMRES can be bounded from below by three factors, Jan 1999

A411 Juhani Pitkäranta
The first locking-free plane-elastic finite element: historia mathematica, Jan 1999

A410 Kari Eloranta
Bounded Triangular and Kagomé Ice, Jan 1999

HELSINKI UNIVERSITY OF TECHNOLOGY INSTITUTE OF MATHEMATICS
RESEARCH REPORTS

The list of reports is continued inside. Electronical versions of the reports are available at *http://www.math.hut.fi/reports/* .

A427 Marko Huhtanen
A Hermitian Lanczos method for normal matrices, Jan 2000

A426 Timo Eirola
A Refined Polar Decomposiotion: $A = UPD$, Jan 2000

A425 Hermann Brunner, Arvet Pedas and Gennadi Vainikko
Piecewise polynomial collocation methods for linear Volterra integro-differential equations with weakly singular kernels, Dec 1999

A424 Olavi Nevanlinna
Resolvent conditions and powers of operators, Dec 1999

A423 Jukka Liukkonen
Generalized Electromagnetic Scattering in a Complex Geometry, Nov 1999

ISBN 951-22-4847-6
ISSN 0784-3143
Intitute of Mathematics, Espoo, 2000