International Journal of Pure and Applied Mathematics
Volume 47 No. 4 2008, 501-518

ITERATIVE REFINEMENT FOR INVARIANT SUBSPACES OF MATRICES WITH APPLICATION TO THE PROMETHEE-GAIA METHOD

Mario Ahues[1 §], David Duvivier[2], Alain Largillier[3], Nadine Meskens[4]
[1,3]Laboratoire de Mathématiques de l'Université de Saint-Etienne (LaMUSE)
Université Jean Monnet
23 Rue du Dr. Paul Michelon
Saint-Étienne, 42100, FRANCE
[1]e-mail: mario.ahues@univ-st-etienne.fr
[3]e-mail: larg@univ-st-etienne.fr
[2]Laboratoire d'Informatique du Littoral
Université du Littoral Côte d'Opale
Maison de la Recherche Blaise Pascal
50, Rue Ferdinand Buisson - BP 719, Calais, 62228, FRANCE
e-mail: david.duvivier@lil.univ-littoral.fr
[4]Louvain School of Management and Catholic
University of Mons (FUCaM)
151, Chaussée de Binche, Mons, 7000, BELGIUM
e-mail: nadine.meskens@fucam.ac.be

Abstract: Recent results in operational research have provided decision makers with an adaptable tool that is able to synthesize the performance measures of several solutions to be compared in a short time. In the context of multi-criteria limited-time decision making problems, two of the authors have developed hybrid models composed of two mathematical models, a set of dedicated heuristics, a stochastic local search, meta-heuristic and a simulation model. According to the decision makers, the solutions are ranked on the basis of sev-

Received: August 18, 2008 © 2008, Academic Publications Ltd.

[§]Correspondence author

eral criteria whose importance determines this ranking. A multicriteria method is incoporated into the hybrid one. In order to summarize the huge amount of resulting data/information, we have embeded the Promethee II multicriteria method and the Gaia plane. This extension requires to compute eigenelements on the output produced by Promethee what is done in two steps: 1) Get rough approximations to the desired part of the spectrum and its spectral subspace; 2) Refine them with an iterative scheme. A Newton-based scheme is proposed and applied to a matrix issued from the Promethee-Gaia method.

AMS Subject Classification: 32P05
Key Words: operational research, hybrid models, Promethee II multicriteria method, Gaia plane, Promethee-Gaia method

1. The General Framework

Let $\mathbb{C}^{n\times n}$ (resp. $\mathbb{R}^{n\times n}$) denote the complex (resp. real) algebra of square matrices of order n with complex (resp. real) entries. The identity matrix of order n will be denoted by I or I_n if the explicit identification of the order is important, all null matrices by O, the spectrum by sp, the resolvent set by re, the spectral radius by ρ and the trace by tr. The matrix norms under consideration are the following: For any matrix M,

$$\begin{aligned} \|\mathsf{M}\|_{\mathrm{F}} &:= \Big(\sum_{i,j} |\mathsf{M}(i,j)|^2\Big)^{\frac{1}{2}}, \\ \|\mathsf{M}\|_2 &:= \max\{\|\mathsf{M}\mathsf{x}\|_{\mathrm{F}} \ : \ \|\mathsf{x}\|_{\mathrm{F}} = 1\}, \end{aligned}$$

where $|\mathsf{M}(i,j)|$ denotes the absolute value of M's entry in row i and column j. The adjoint of M will be denoted by M^*.

Proposition 1. *Let M be any matrix.*

1. *The norms $\|\cdot\|_{\mathrm{F}}$ and $\|\cdot\|_2$ satisfy:*
   - *(a)* $\|\mathsf{M}\|_{\mathrm{F}}^2 = \mathrm{tr}(\mathsf{M}^*\mathsf{M})$,
   - *(b)* $\|\mathsf{M}\|_2^2 = \rho(\mathsf{M}^*\mathsf{M})$, *and*
   - *(c)* $\|\mathsf{M}\|_2 \le \|\mathsf{M}\|_{\mathrm{F}}$.
2. *For any matrix N of appropriate size,*
   - *(a)* $\|\mathsf{M}\mathsf{N}\|_2 \le \|\mathsf{M}\|_2\|\mathsf{N}\|_2$, *and*

(b) $\|\mathsf{MN}\|_{\mathrm{F}} \leq \min\{\|\mathsf{M}\|_2\|\mathsf{N}\|_{\mathrm{F}}, \|\mathsf{N}\|_2\|\mathsf{M}\|_{\mathrm{F}}\}$.

3. For any unitary matrix Q *of appropriate size,*

(a) $\|\mathsf{QM}\|_{\mathrm{F}} = \|\mathsf{M}\|_{\mathrm{F}}$, *and*
(b) $\|\mathsf{QM}\|_2 = \|\mathsf{M}\|_2$.

4. For any unitary matrix Q *of appropriate size,*

(a) $\|\mathsf{MQ}\|_{\mathrm{F}} = \|\mathsf{M}\|_{\mathrm{F}}$, *and*
(b) $\|\mathsf{MQ}\|_2 = \|\mathsf{M}\|_2$.

The spectral problem will be formulated using operators defined in the space $\mathbb{C}^{n\times m}$ where $m \leq n$. The norm of a linear operator $L : \mathbb{C}^{n\times m} \to \mathbb{C}^{n\times m}$ will be

$$\|L\| := \max\{\|L(\mathsf{X})\|_2 \ : \ \mathsf{X} \in \mathbb{C}^{n\times m},\ \|\mathsf{X}\|_2 = 1\}.$$

Let $\mathsf{A} \in \mathbb{C}^{n\times n}$ be a nonsingular matrix. Let us suppose that with the help of a stable numerical algorithm, we have computed a matrix $\mathsf{X}_0 \in \mathbb{C}^{n\times m}$ whose columns are the canonical coordinates of an approximation to a basis for some maximal invariant subspace of A. Let Λ denote the corresponding spectral set. In the same way, let the columns of $\mathsf{Y}_0 \in \mathbb{C}^{n\times m}$ be the canonical coordinates of an approximation to a basis for the maximal invariant subspace of A^* corresponding to $\overline{\Lambda}$, and assume that bases X_0 and Y_0 are normalized by $\mathsf{Y}_0^*\mathsf{X}_0 = \mathsf{X}_0^*\mathsf{X}_0 = \mathsf{I}_m$. In other words, X_0 represents an orthonormal basis, and Y_0 is biorthogonal to X_0.

The unit error of the finite precision arithmetic device will be denoted by $\mathbf{u}$. The stability of the algorithm implies (see [8]) the existence of matrices A_0 and E such that:

1. X_0 is an exact basis of the maximal invariant subspace of A_0 corresponding to a spectral set Λ_0 of A_0.

2. Y_0 is an exact basis of the maximal invariant subspace of A_0^* corresponding to the conjugate set $\overline{\Lambda_0}$.

3. $\mathsf{A}_0 = \mathsf{A} + \mathsf{E}$.

4. For some $c > 0$ independent of (A, E): $\|\mathsf{E}\|_2 \leq c\,\mathbf{u}\,\|\mathsf{A}\|_2$.

Define $\Theta_0 := \mathsf{Y}_0^*\mathsf{A}_0\mathsf{X}_0$.

Proposition 2. *The following results hold:*

1. The spectral projection of A_0 *corresponding to its spectral set* Λ_0 *has as canonical matrix representation*

$$\mathsf{P}_0 := \mathsf{X}_0\mathsf{Y}_0^*.$$

2. Θ_0 *represents in the basis* X_0 *the restriction to the maximal invariant subspace of the linear map canonically represented by the matrix* A_0.

3. The Sylvester matrix operator $G_0 : \mathbb{C}^{n\times m} \to \mathbb{C}^{n\times m}$ *defined by*

$$G_0(\mathsf{X}) := (\mathsf{I} - \mathsf{P}_0)\mathsf{A}_0\mathsf{X} - \mathsf{X}\Theta_0 \quad \textit{for all } \mathsf{X} \in \mathbb{C}^{n\times m},$$

is nonsingular.

Recall (see [4]) that the bloc-reduced resolvent operator of A_0 corresponding to its spectral set Λ_0 is defined to be

$$\Sigma_0(\mathsf{X}) := G_0^{-1}[(\mathsf{I} - \mathsf{P}_0)\mathsf{X}] \quad \text{for all } \mathsf{X} \in \mathbb{C}^{n\times m}.$$

2. Newton-Kantorovich

The elements of Λ_0 will play the role of approximations to the eigenvalues of A contained in the set Λ. If X_∞ represents a basis of the exact maximal invariant subspace of A corresponding to Λ, then each column of AX_∞ is a linear combination of the columns of X_∞, i.e. there exists a matrix $\Theta \in \mathbb{C}^{m\times m}$ such that $\mathsf{AX}_\infty = \mathsf{X}_\infty\Theta$. If Y is any adjoint to X_∞, then $\Theta = \mathsf{Y}^*\mathsf{AX}_\infty$. If P_0 is sufficiently close to the exact spectral projection, then there exists an exact basis representation X_∞ such that $\mathsf{Y}_0^*\mathsf{X}_\infty = \mathsf{I}_m$. Hence, X_∞ is a root of the quadratic Sylvester opertor F defined by

$$F(\mathsf{X}) := \mathsf{AX} - \mathsf{XY}_0^*\mathsf{AX} \quad \text{for all } \mathsf{X} \in \mathbb{C}^{n\times m}.$$

The root X_∞ can thus be found by a Newton-Kantorovich iterative process starting with X_0. The Fréchet derivative of F at a given matrix $\mathsf{X} \in \mathbb{C}^{n\times m}$ is the linear Sylvester operator

$$F'(\mathsf{X})(\mathsf{H}) := \mathsf{AH} - \mathsf{HY}_0^*\mathsf{AX} - \mathsf{XY}_0^*\mathsf{AH} \quad \text{for all } \mathsf{H} \in \mathbb{C}^{n\times m}.$$

Using the current iterate X_{cur} and if $F'(\mathsf{X}_{\text{cur}})$ is bijective, Newton's method computes the next iterate as the unique matrix $\mathsf{X}_{\text{new}} \in \mathbb{C}^{n\times m}$ satisfying the linear Sylvester equation

$$F'(\mathsf{X}_{\text{cur}})(\mathsf{X}_{\text{new}} - \mathsf{X}_{\text{cur}}) = -F(\mathsf{X}_{\text{cur}}).$$

In our case, this equation can be rewritten as

$$(\mathsf{I} - \mathsf{X}_{\text{cur}}\mathsf{Y}_0^*)\mathsf{AX}_{\text{new}} - \mathsf{X}_{\text{new}}\Theta_{\text{cur}} = -\mathsf{X}_{\text{cur}}\Theta_{\text{cur}}, \tag{1}$$

where $\Theta_{\text{cur}} := \mathsf{Y}_0^*\mathsf{AX}_{\text{cur}} \in \mathbb{C}^{m\times m}$.

Proposition 3. *The following results hold:*

1. Each iterate X_{cur} *satisfies* $\mathsf{P}_0\mathsf{X}_{\text{cur}} = \mathsf{X}_0$.

2. Each residual $F(\mathsf{X}_{\rm cur})$ *satisfies* $\mathsf{P}_0 F(\mathsf{X}_{\rm cur}) = \mathsf{O}$.

For our purposes, the Kantorovich a posteriori convergence theorem reads as follows.

Theorem 1. *Suppose that* F, X_0, $c_0 > 0$, $\ell > 0$ *and* $m_0 > 0$ *satisfy:*

1. $F'(\mathsf{X}_0)$ *is bijective and* $\|F'(\mathsf{X}_0)^{-1}\| \le m_0$,

2. $\|F'(\mathsf{X}_0)^{-1}(F(\mathsf{X}_0))\| \le c_0$,

3. ℓ *is a Lipschitz constant for* F' *on*

$$\mathcal{D}_0 := \{\mathsf{X} \in \mathbb{C}^{n\times m} \ : \ \|\mathsf{X} - \mathsf{X}_0\| \le 2c_0\},$$

4. $m_0 \ell c_0 < 1/2$.

Then, F *has a unique zero* X_∞ *in* $\mathcal{D}_0$ *and Newton's iterative process generates a well defined sequence* X_k *such that, for all* $k \ge 0$,

$$\|\mathsf{X}_{k+1} - \mathsf{X}_\infty\| \le \frac{m_0 \ell}{1 - 2m_0 \ell c_0} \|\mathsf{X}_k - \mathsf{X}_\infty\|^2,$$

i.e. convergence is quadratic.

Proof. See [1]. □

The following theorem shows that Kantorovich hypotheses hold for a sufficiently stable algorithm.

Theorem 2. *Suppose that the perturbation* E *introduced by the algorithm used in the computation of* X_0 *and* Y_0 *is such that*

$$\|\mathsf{E}\|_2 < \frac{\nu}{2\,\|G_0^{-1}\|},$$

where

$$\nu := \min\left\{ \frac{1}{1 + 2\,\|\mathsf{Y}_0\|_2}, \frac{1}{8\,\|G_0^{-1}\|\,\|\mathsf{Y}_0\|\,\|\mathsf{A}\|_2\,(1 + \|\mathsf{Y}_0\|_2)} \right\}.$$

Then

$$m_0 := 2\,\|G_0^{-1}\|, \quad \ell := 2\,\|\mathsf{A}\|_2\,\|\mathsf{Y}_0\|_2,$$

and

$$c_0 := 2\,\|G_0^{-1}\|\,\|\mathsf{E}\|_2 (1 + \|\mathsf{Y}_0\|_2)$$

satisfy the hypotheses of Theorem 1.

Proof. Since

$$F(\mathsf{X}_0) = \mathsf{A}\mathsf{X}_0 - \mathsf{X}_0\mathsf{Y}_0^*\mathsf{A}\mathsf{X}_0 = \mathsf{X}_0\mathsf{Y}_0^*\mathsf{E}\mathsf{X}_0 - \mathsf{E}\mathsf{X}_0,$$

and since $\|\mathsf{X}_0\|_2 = 1$ and $\|\mathsf{Y}_0^*\|_2 = \|\mathsf{Y}_0\|_2$, we get

$$\|F(\mathsf{X}_0)\|_2 \leq \|\mathsf{E}\|_2(1 + \|\mathsf{Y}_0\|_2).$$

On the other hand,

$$F'(\mathsf{X}_0)(\mathsf{H}) = (G_0 - E_0)(\mathsf{H}),$$

where

$$E_0(\mathsf{H}) := (\mathsf{I} - \mathsf{P}_0)\mathsf{E}\mathsf{H} - \mathsf{H}\mathsf{Y}_0^*\mathsf{E}\mathsf{X}_0.$$

Writing

$$F'(\mathsf{X}_0) = G_0 - E_0 = G_0(I - G_0^{-1}E_0)$$

we conclude that, if

$$\|E_0\| < \frac{1}{\|G_0^{-1}\|} \tag{2}$$

then $F'(\mathsf{X}_0)$ is bijective and

$$\|F'(\mathsf{X}_0)^{-1}\| \leq \frac{\|G_0^{-1}\|}{1 - \|G_0^{-1}\|\|E_0\|}.$$

But

$$\|E_0(\mathsf{H})\|_2 \leq (1 + 2\|\mathsf{Y}_0\|_2)\|\mathsf{E}\|_2 \|\mathsf{H}\|_2,$$

hence

$$\|E_0\| \leq (1 + 2\|\mathsf{Y}_0\|_2)\|\mathsf{E}\|_2$$

and

$$\|\mathsf{E}\|_2 < \frac{1}{2\|G_0^{-1}\|(1 + 2\|\mathsf{Y}_0\|_2)}$$

is a sufficient condition for (2) which implies that

$$\|F'(\mathsf{X}_0)^{-1}\| \leq \frac{\|G_0^{-1}\|}{1 - \|G_0^{-1}\|\|E_0\|} \leq m_0 := 2\|G_0^{-1}\|.$$

Also

$$\|F'(\mathsf{X}_0)^{-1}(F(\mathsf{X}_0))\|_2 \leq m_0 \|F(\mathsf{X}_0)\|_2 \leq c_0\,,$$

where

$$c_0 := 2\|G_0^{-1}\| \|\mathsf{E}\|_2(1 + \|\mathsf{Y}_0\|_2).$$

To obtain a Lispschitz constant for F' remark that, for all X, Y, $\mathsf{H} \in \mathbb{C}^{n\times m}$,

$$\|(F'(\mathsf{X}) - F'(\mathsf{Y}))(\mathsf{H})\|_2 \leq 2\|\mathsf{Y}_0\|_2\|\mathsf{A}\|_2\|(\mathsf{Y} - \mathsf{X})\|_2\|\mathsf{H}\|_2.$$

We thus set $\ell := 2\|\mathsf{Y}_0\|_2\|\mathsf{A}\|_2$ and the proof is complete. □

3. Sylvester Equations

The mathematical objects to be computed are Λ_0, X_0, Y_0 and, for a given X_{cur}, the unique matrix X verifying

$$\mathsf{MX} - \mathsf{XN} = \mathsf{S},$$

where $\mathsf{M} := (\mathsf{I} - \mathsf{X}_{\text{cur}}\mathsf{Y}_0^*)\mathsf{A}$, $\mathsf{N} := \Theta_{\text{cur}} = \mathsf{Y}_0^*\mathsf{A}\mathsf{X}_{\text{cur}}$, and $\mathsf{S} := -\mathsf{X}_{\text{cur}}\mathsf{N}$. The resolution of (1) can be performed either with the Bartels-Stewart algorithm (see [6]) or with the Hessenberg-Schur algorithm (see [2], [9] and [7]).

3.1. The Bartels-Stewart Algorithm

Equation (1) can be rewritten in the unknown

$$\vec{\mathsf{X}} := \begin{bmatrix} \mathsf{X}(*,1) \\ \vdots \\ \mathsf{X}(*,m) \end{bmatrix} \in \mathbb{C}^{nm\times 1}$$

as

$$(\mathsf{I}_m \otimes \mathsf{M} - \mathsf{N}^\top \otimes \mathsf{I}_n)\vec{\mathsf{X}} = \vec{\mathsf{S}},$$

and this standard linear system with nm equations and nm unknowns can be solved using a convenient classical linear algebra algorithm, such as Gauss elimination, LU factorization, or a convergent iterative method.

We recall that the Kronecker product $\otimes$ is defined as follows:

For any matrices $\mathsf{A} \in \mathbb{C}^{p\times q}$ and $\mathsf{B} \in \mathbb{C}^{r\times s}$,

$$\mathsf{A}\otimes\mathsf{B} := \begin{bmatrix} \mathsf{A}(1,1)\mathsf{B} & \mathsf{A}(1,2)\mathsf{B} & \cdots & \mathsf{A}(1,q)\mathsf{B} \\ \mathsf{A}(2,1)\mathsf{B} & \mathsf{A}(2,2)\mathsf{B} & \cdots & \mathsf{A}(2,q)\mathsf{B} \\ \vdots & \vdots & \ddots & \vdots \\ \mathsf{A}(p,1)\mathsf{B} & \mathsf{A}(p,2)\mathsf{B} & \cdots & \mathsf{A}(p,q)\mathsf{B} \end{bmatrix} \in \mathbb{C}^{pr\times qs}.$$

3.2. The Hessenberg-Schur Algorithm

The principle is to make (1) equivalent to m standard linear systems of order n. For this purpose, let U be a unitary matrix triangularizing N and V a unitary matrix producing a Hessenberg form of M. Then (1) can be rewritten as

$$\mathsf{HY} - \mathsf{Y}\Delta = \mathsf{K},$$

where:

$$\begin{array}{rcll} \Delta & := & \mathsf{U}^*\mathsf{N}\mathsf{U} & \text{is an upper triangular matrix in } \mathbb{C}^{m\times m}, \\ \mathsf{H} & := & \mathsf{V}^*\mathsf{M}\mathsf{V} & \text{is an upper Hessenberg matrix in } \mathbb{C}^{n\times n}, \\ \mathsf{K} & := & \mathsf{V}^*\mathsf{S}\mathsf{U} & \text{is the new second member in } \mathbb{C}^{n\times m}, \\ \mathsf{Y} & := & \mathsf{V}^*\mathsf{X}\mathsf{U} & \text{is the new unknown in } \mathbb{C}^{n\times m}. \end{array}$$

This new Sylvester equation is equivalent to the following m linear systems:

$$(\mathsf{H}-\Delta(1,1)\mathsf{I}_n)\mathsf{Y}(*,1)=\mathsf{K}(*,1), \text{ and for } j\in[\![2,m]\!]:$$
$$(\mathsf{H}-\Delta(j,j)\mathsf{I}_n)\mathsf{Y}(*,j)=\mathsf{K}(*,j)+\sum_{i=1}^{j-1}\Delta(i,j)\mathsf{Y}(*,i).$$

The fact that H is an upper Hessenberg matrix is not essential but makes easy the Gauss elimination procedure or the LU factorization. The major drawback of this procedure is that the computation of the Schur unitarily similar matrix Δ cannot be done in a finite number of steps. That is why it may be advisable to complete the algorithm with an iterative refinement scheme as follows. Let $\widehat{\Delta} := \widehat{\mathsf{U}}^*\mathsf{N}\widehat{\mathsf{U}}$ be the computed version of Δ, where $\widehat{\mathsf{U}}$ is a unitary matrix. We assume that $\widehat{\Delta}$ and N are indeed similar matrices. Let T be the upper triangular part of $\widehat{\Delta}$. We suppose that $\widehat{\Delta}$ is quasi-uppertriangular in the sens that in some matrix norm, $\widehat{\Delta}-\mathsf{T}$ is sufficiently small and in any case small enough to make convergent the following iterative process:

$k \leftarrow 0$
Compute $\mathsf{Y}^{(k)}$ such that $\mathsf{H}\mathsf{Y}^{(k)}-\mathsf{Y}^{(k)}\mathsf{T}=\mathsf{K}$
Compute the Residual $\mathsf{R}^{(k)} \leftarrow \mathsf{H}\mathsf{Y}^{(k)}-\mathsf{Y}^{(k)}\widehat{\Delta}-\mathsf{K}$
While Stop Criterion Not Satisfied Do
$k \leftarrow k+1$
Compute $\mathsf{Y}^{(k)}$ such that
$\quad \mathsf{H}\mathsf{Y}^{(k)}-\mathsf{Y}^{(k)}\mathsf{T}=\mathsf{Y}^{(k-1)}(\widehat{\Delta}-\mathsf{T})+\mathsf{K}$
Compute the Residual $\mathsf{R}^{(k)} \leftarrow \mathsf{H}\mathsf{Y}^{(k)}-\mathsf{Y}^{(k)}\widehat{\Delta}-\mathsf{K}$
Done

For clusters of eigenvalues with total algebraic mutiplicity $m \leq 3$ we have chosen to solve (1) with the Bartels-Stewart algorithm and the LU factorization. The Hessenberg-Schur algorithm coupled with the refinement process is applied in all other cases.

4. Computing the Starting Point of Iterations

We consider le QR Francis method for the Schur form of a general matrix. The basic algorithm reads as:

$k \leftarrow 0$
$\mathsf{A}_0 \leftarrow \mathsf{A}$
While Stop Criterion Not Satisfied Do
$k \leftarrow k+1$
$\mathsf{Q}_k \leftarrow$ a unitary left factor of A_{k-1}
$\mathsf{R}_k \leftarrow$ the upper-triangular corresponding right factor
$\mathsf{A}_k \leftarrow \mathsf{R}_k\mathsf{Q}_k$
Done

A well-known result concerning the convergence of this algorithm is the following:

Theorem 3. *The QR method is an implementation of the Krylov method in a nested way: For all $i, j \in [\![1, n]\!]$ and $k \geq 1$, $S_{j,k} := \mathsf{A}^k(S_{j,0})$, $\dim S_{j,0} = j$, $S_{i,0} \subsetneq S_{j,0}$ for $i < j$. The Krylov method converges if:*

(H1) $|\lambda_1| > |\lambda_2| > \ldots > |\lambda_n| > 0$, *and*

(H2) For all $i \in [\![1, n]\!]$, $\dim P^{(i)}(S_{i,0})) = i$, where $P^{(i)}$ is the spectral projection corresponding to the spectral set $\{\lambda_j \ : \ 1 \leq j \leq i\}$.

In order to satisfy condition *(H1)* we can introduce heuristic shifts σ_k:

$k \leftarrow 0$
$\mathsf{A}_k \leftarrow \mathsf{A}$
While Stop Criterion Not Satisfied Do
$k \leftarrow k+1$
$\mathsf{Q}_k \leftarrow$ a unitary left factor of $\mathsf{A}_{k-1} - \sigma_{k-1}\mathsf{I}$
$\mathsf{R}_k \leftarrow$ the upper-triangular corresponding right factor
$\mathsf{A}_k \leftarrow \mathsf{R}_k\mathsf{Q}_k + \sigma_{k-1}\mathsf{I}$
Done

Condition (H2) is automaticaly satisfied if A is an unreduced Hessenberg matrix

$$\begin{aligned} \mathsf{A}(i, i-1) &\neq 0 \text{ for } i \in [\![2, n]\!], \\ \mathsf{A}(i, j) &= 0 \text{ for } j \in [\![1, n-2]\!] \text{ and } i \in [\![j+2, n]\!], \end{aligned}$$

and if the starting subspace $S_{j,0}$ is spanned by the first j canonical vectors for each $j \in [\![1, n]\!]$.

Concerning the stopping criterion, we suppose that A is already in an upper

unreduced Hessenberg form:

$$\mathsf{A}(i,j) \;=\; 0 \quad \text{for all } i,j \in [\![1,n]\!] \text{ such that } i-j \geq 2$$

and that it satisfies (H1). Let A_k be the k-th iterate of the basic QR algorithm. Then $\mathsf{A}_k = \mathsf{T}_k + \mathsf{E}_k$, where T_k is the upper triangular part of A_k and

$$\mathsf{E}_k \;=\; \sum_{j=1}^{n-1} \mathsf{A}_k(j+1,j)\mathsf{e}_j\mathsf{e}_{j+1}^*,$$

e_i being the i-th canonical column in $\mathbb{C}^{n\times 1}$. Under the given hypotheses,

$$\begin{aligned} \lim_{k\to\infty} \mathsf{E}_k &= \mathsf{O}, \\ \lim_{k\to\infty} \mathsf{A}_k(i,i) &= \lambda_i(\mathsf{A}) \quad \text{for all } i \in [\![1,n]\!]. \end{aligned}$$

Hence, for all $\epsilon \in]0,1[$ there exists k_ϵ such that, for $k > k_\epsilon$,

$$\sum_{j=1}^{n-1} |\mathsf{A}_k(j+1,j)| \;<\; \epsilon.$$

From Rayleigh-Schrödinger theory, there exists $\epsilon_0 \in]0,1[$ such that, for all $\epsilon \in]0,\epsilon_0[$ and all $k > k_\epsilon$, the eigenvalues of

$$\mathsf{H}_k(t) := \mathsf{T}_k + t\,\mathsf{E}_k,$$

have a power series expansion in t around the eigenvalues of T_k which is convergent for all $t \in [0,1]$. Concerning the eigenvalue expansion, it is shown in [5] that

$$\lambda_i(\mathsf{A}) \;=\; \mathsf{A}_k(i,i) + \mu_i + O(\epsilon^2),$$

where

$$\begin{aligned} \mu_1 &= \frac{\mathsf{A}_k(1,2)\mathsf{A}_k(2,1)}{\mathsf{A}_k(1,1) - \mathsf{A}_k(2,2)}; \;\; \text{for } i \in [\![2,n-1]\!]: \\ \mu_i &= \frac{\mathsf{A}_k(i,i-1)\mathsf{A}_k(i-1,i)}{\mathsf{A}_k(i,i) - \mathsf{A}_k(i-1,i-1)} + \frac{\mathsf{A}_k(i,i+1)\mathsf{A}_k(i+1,i)}{\mathsf{A}_k(i,i) - \mathsf{A}_k(i+1,i+1)}; \\ \mu_n &= \frac{\mathsf{A}_k(n,n-1)\mathsf{A}_k(n-1,n)}{\mathsf{A}_k(n,n) - \mathsf{A}_k(n-1,n-1)}. \end{aligned}$$

In order to reduce the size of the problem by partitioning matrix A_k into 2 blocks because the entry $\mathsf{A}_k(\ell+1,\ell)$ is found to be negligible, we propose the joint verification of:

$$|\mathsf{A}_k(\ell+1,\ell)| < \epsilon, \quad |\mu_\ell| < \epsilon |\mathsf{A}_k(\ell,\ell)|.$$

In order to stop iterations, we suggest – taking into account these remarks – that the stop criterion contains at least the following requirements, as a mean

to get approximate eigenvalues with a relative precision of the order of ϵ:

$$\begin{aligned} \sum_{j=1}^{n-1} |\mathsf{A}_k(j+1,j)| &< \epsilon_*, \\ \forall j \in [\![1,n]\!], \quad |\mu_j| &< \epsilon |\mathsf{A}_k(j,j)|, \end{aligned}$$

where ϵ_* is sufficiently small to ensure the convergence of the Rayleigh-Schrödinger series and $\epsilon \ll \epsilon_*$. A typical choice is to take ϵ_* of the order of the simple precision and ϵ of the order of the double precision.

In the case of an Hermitian data we may assume that A is its reduction to a unitarily similar real symmetric tridiagonal matrix. In that case, it is shown in [3] that

$$\lambda_i(\mathsf{A}) = \mathsf{A}_k(i,i) + \mu_i + O(\epsilon^4),$$

where

$$\begin{aligned} \mu_1 &= \frac{\mathsf{A}_k(1,2)^2}{\mathsf{A}_k(1,1) - \mathsf{A}_k(2,2)}; \text{ for } i \in [\![2, n-1]\!]: \\ \mu_i &= \frac{\mathsf{A}_k(i,i-1)^2}{\mathsf{A}_k(i,i) - \mathsf{A}_k(i-1,i-1)} + \frac{\mathsf{A}_k(i,i+1)^2}{\mathsf{A}_k(i,i) - \mathsf{A}_k(i+1,i+1)}; \\ \mu_n &= \frac{\mathsf{A}_k(n,n-1)^2}{\mathsf{A}_k(n,n) - \mathsf{A}_k(n-1,n-1)}. \end{aligned}$$

Hence the suggested stop criterion will give approximate eigenvalues with a relative precision of the order of ϵ^4.

The starting basis of the iterative refinement process can be computed with the simultaneous inverse iteration algorithm coupled with an orthonormalization step. For an invariant subspace with dimension m, the algorithm reads:

$k \leftarrow 0$
Y_k full column rank matrix randomly chosen in $\mathbb{C}^{n\times m}$
$\mathsf{U}_k \leftarrow$ a unitary left factor of Y_k
While Stop Criterion Not Satisfied Do
$k \leftarrow k+1$
Compute Y_k such that $(\mathsf{A} - \widehat{\sigma}\mathsf{I})\mathsf{Y}_k = \mathsf{U}_{k-1}$
$\mathsf{U}_k \leftarrow$ a unitary left factor of Y_k
Done

where $\widehat{\sigma}$ is the arithmetic mean, counting algebraic multiplicities, of the approximate eigenvalues computed before and corresponding to the invariant subspace under consideration. The residual criterion is based on the fact that, for an exact orthonormal basis U of an invariant subspace of A, $\mathsf{AU} = \mathsf{UU}^*\mathsf{AU}$. Hence, for some computable matrix norm $\|\cdot\|$, and a given relative precision ϵ, the

residual criterion checks if the iterate U_k is such that $\|(\mathsf{I}-\mathsf{U}_k\mathsf{U}_k^*)\mathsf{A}\mathsf{U}_k\| < \epsilon$. The simultaneous inverse iteration with orthonormalization is the power method applied to the matrix $(\mathsf{A}-\widehat{\sigma}\mathsf{I})^{-1}$. It converges linearly under the condition that $\widehat{\sigma}$ is well separated from the other weighted means of approximate eigenvalues and provided that the starting randomly chosen basis Y_0 has a full rank spectral projection onto the exact invariant subspace. For more details, see [4] and [8].

5. To Cluster or Not to Cluster: An Example

Let

$$\mathsf{A} := \begin{bmatrix} 1.000 & 0.005 & 0.002 & 0.002 \\ 0.005 & 1.000 & 0.005 & 0.005 \\ 0.002 & 0.005 & 2.000 & 0.005 \\ 0.002 & 0.005 & 0.005 & 2.000 \end{bmatrix}.$$

The QR algorithm recognizes 3 clusters of approximate eigenvalues which are a spectral set $\widetilde{\Lambda}_0(\mathsf{A})$ with algebraic multiplicity 2, and two simple eigenvalues $\widetilde{\lambda}_1(\mathsf{A})$ and $\widetilde{\lambda}_2(\mathsf{A})$. The iterative refinement computations for cluster $\widetilde{\Lambda}_0(\mathsf{A})$ stops after 2 iterations, for cluster $\widetilde{\lambda}_1(\mathsf{A})$ after 6 iterations, and for cluster $\widetilde{\lambda}_2(\mathsf{A})$ after 5 iterations. Residuals for the iteratively refined approximate bases decay as shown in Table 1. Clustering is thus strongly recommended.

Iteration	$\Lambda_0(\mathsf{A})$	$\lambda_1(\mathsf{A})$	$\lambda_2(\mathsf{A})$
1	$1.28E-03$	$4.76E-03$	$4.98E-03$
2	$1.99E-09$	$2.91E-03$	$2.68E-05$
3	$1.12E-16$	$2.29E-04$	$2.43E-08$
4		$1.99E-06$	$2.00E-14$
5		$1.57E-10$	$1.12E-16$
6		$2.23E-16$	

Table 1

6. Application to the Promethee-Gaia Method

In a wide variety of applications, the decision makers hesitate between several solutions. This observation led to the introduction of a preference model including two thresholds: the indifference threshold, under which the decision makers

are indifferent, and the preference threshold, above which the decision makers have an indubitable strict preference. The intermediate area is considered as a region, where the decision makers hesitate between preference and indifference.

Decision makers need a highly adaptable tool which must be able to quickly synthesize the performance measures of several solutions to be compared in a limited time. Consequently, the resulting tool must incorporate gradation, tinge and fuzziness in the judgment of decision makers while comparing several solutions. Moreover, it is necessary to replace the traditional concept of an optimal solution by a concept of a *satisfactory solution*, according to decision makers' point of view. In the context of multicriteria limited-time decision making problems, we have developed several hybrid models composed of two mathematical models, a set of dedicated heuristics, a stochastic local search, meta-heuristic and a simulation model. According to the decision makers, the solutions are ranked on the basis of several criteria computed either by the simulation model or the decision makers. The relative importance of these criteria, defined by the decision makers, determines this ranking. But this is neither easy nor realistic when considering some real-life problems since when comparing two solutions, a decision maker often accepts a solution which is worse from the point of view of the main criterion, if this solution leads to significant improvements on some other criteria. The best way to achieve the goal is to incorporate a multicriteria method into the hybrid itself (see [12], [13], [14], [16] and the references therein).

One important issue when comparing several solutions is to summarize the huge amount of resulting data/information so as to provide a synthetic and accurate view of the process to the decision makers. That is one of the reasons why we choose to embed the Promethee II multicriteria method (*Preference Ranking Organization Method for Enrichment Evaluations*, see [10], [11] and [15]) and one of its extensions, the Gaia plane (*Geometrical Analysis for Interactive Assistance*). This extension requires to compute eigenelements on the results given by the Promethee method.

Promethee II provides a complete ranking based on pairwise comparisons of solutions and permits to simultaneously maximize a set of criteria and minimize another set. In the following paragraphs, a and b (and later $a_1, a_2, \ldots, a_n$) denote potential alternatives, i.e. solutions. Promethee builds an outranking relation using a preference function for each criterion, which represents the decision makers' preference $P_j(a, b)$ for a solution a with regard to a solution b on a given criterion j. $P_j(a, b) = 0$ means indifferent choice between solution a and solution b, there is no preference for solution a over solution b, $P_j(a, b)$

close to zero means that there is a weak preference for solution a over solution b, $P_j(a,b)$ close to 1 means that there is a strong preference for solution a over solution b and $P(a,b)=1$ means that there is a strict preference for solution a over solution b.

The decision makers have to give additional information on each criterion: the preference function P_j and the weight ω_j for criterion j. The outranking index of the solution a over the solution b is

$$\pi(a,b)=\frac{\sum_{j=1}^{k}\omega_j P_j(a,b)}{\sum_{j=1}^{k}\omega_j},$$

where k is the number of criteria. This index provides a measure of the preference for a on b over all the criteria. On this basis, Promethee II computes positive and negative preference flows for each solution and

$$\phi^+(a)=\frac{1}{n-1}\sum_{x=1}^{n}\pi(a,x), \qquad \phi^-(a)=\frac{1}{n-1}\sum_{x=1}^{n}\pi(x,a),$$
$$\phi(a)=\phi^+(a)-\phi^-(a).$$

The out-flow (or positive preference flow) expresses how much a solution is dominating the other solutions, and the in-flow (or negative preference flow) how much it is dominated by the other solutions. Based on the net outranking flows, the selected version labelled Promethee II provides a total order of the solutions, see [11]. The solution a outranks the solution b if, and only if, $\phi(a) > \phi(b)$, and solutions a and b are indifferent solutions if, and only if, $\phi(a)=\phi(b)$.

The information relative to a decision problem including k criteria is represented in a k-dimensional space. The whole information represents a huge quantity of criteria and solutions to be analyzed in a limited time by the decision makers. So it is necessary to present the results in such a way that the decision makers are able to identify, at first glance, the best solutions and their quality relative to the different criteria. To achieve this goal, a bidimensional representation of the problem named the Gaia plane has been introduced in [10]. In this representation the location of the alternatives with respect to the criteria can be observed. The Gaia plane is obtained by projection of this information on a plane such that as few information as possible is lost. Solutions are represented by points and, criteria by arrows. Criteria represented by arrows with similar directions indicate that they have similar discrimination power with respect to the alternatives. If they appear in opposite directions, they are conflicting. In addition to the solutions and criteria, the projection of the weight vector in the Gaia plane corresponds to another arrow which shows the direction of the compromise resulting from the weights allocated to the criteria

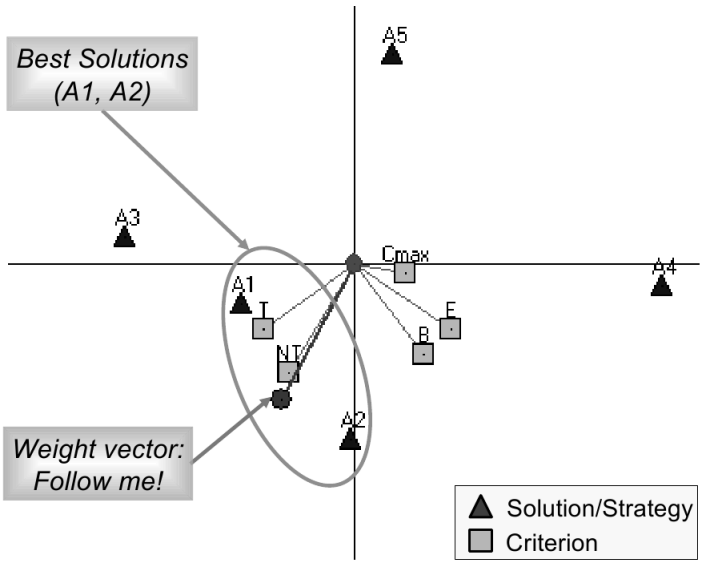

Figure 1:

and represents the objective resulting from weighing the criteria after making them homogeneous. Decision makers are invited to consider solutions located in that direction. The length of this vector indicates its decision power. From a technical point of view the Gaia plane is obtained by applying a PCA (Principal Component Analysis) on the square matrix A defined through the individual net flows by:

$$\mathsf{A}(i,j) := \phi_j(a_i).$$

The individual net flows $\phi_j(a)$ are obtained via the following equations:

$$\phi_j^+(a) = \frac{1}{n-1}\sum_{x=1}^{n} P_j(a,x), \qquad \phi_j^-(a) = \frac{1}{n-1}\sum_{x=1}^{n} P_j(x,a),$$
$$\phi_j(a) = \phi_j^+(a) - \phi_j^-(a).$$

Figure 1 shows that Gaia provides an approximate representation of the information related to this problem, even if up to 92.74% of the variance is reproduced by the first two principal components. In the present example, five scheduling strategies (i.e. five solutions to the studied problem) are compared. They are labeled A1, A2, A3, A4 and A5. These solutions are compared on the basis of five criteria. Listed in decreasing order of importance according to decision makers choices, these criteria are named NT, T, Cmax, E and B. The

Refined eigenvalue	Refined eigenvector
1.103484797260972E+0	3.860206519940642E−1 5.292946853217178E−1 –2.970855085380545E−1 –5.626386981636733E−1 –4.074790951735133E−1
0.537263949030415E+0	–6.403423955551950E−1 –3.875291168467758E−1 –5.583675564808065E−2 –3.859439584728704E−1 –5.363887283269551E−1
0.095596108780426E+0	–3.192920518952169E−1 4.415513843812426E−2 –6.933806260813352E−1 –2.435381995253115E−1 5.966811754996950E−1
0.032886082428202E+0	4.101151273677964E−1 –5.513908622178885E−1 2.585397586455628E−1 –6.019833393549611E−1 3.149986445168979E−1
0.000000000000000E+0	4.146674838644793E−1 –5.152454254481501E−1 –6.026737202496584E−1 3.372246165867986E−1 –3.026816465116050E−1

Table 2:

decision stick reflects these choices, inviting the decision makers to consider the solutions located in that direction. Here, the solutions to be selected are A2 and A1 with a small preference for A2 over A1.

With residuals of the order of 10^{-15}, the refined approximate eigenelements are shown in Table 2.

References

[1] M. Ahues, Newton methods with Hölder derivative, *Numer. Func. Anal. and Optimiz.*, **25**, No-s: 5 and 6 (2004), 1-17.

[2] M. Ahues, L. Grammont, On the numerical analysis of generalized Sylvester equations. Part I: A theoretical spectral framework, *Numer. Func. Anal. and Optimiz.*, **15**, No-s: 7 and 8 (1994), 713-727.

[3] M. Ahues, A. Largillier, Spectral Rayleigh-Schr̈odinger series revisited, *Technical Report*, Catholic University of Mons (FUCaM), Belgium (2007).

[4] M. Ahues, A. Largillier, B.V. Limaye, *Spectral Computations with Bounded Operators*, Chapman and Hall, CRC, Boca Raton (2001).

[5] M. Ahues, F. Tisseur, A new deflation criterion for the QR algorithm, UT CS-97-353, www.netlib.org/lapack/lawnspdf/lawn122.pdf (1997).

[6] R. Bartels, G. Stewart, Solution of the matrix equation $AX + XB = C$: Algorithm 432, *Comm. ACM*, **15** (1972), 820-826.

[7] G.H. Golub, S. Nashed, C.F. Van Loan, A Hessenberg-Schur form for the problem $AX+XB = C$, *IEEE Trans. Autom. Control*, **24** (1979), 909-913.

[8] G.H. Golub, C.F. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, MD, 3-rd Edition (1996).

[9] L. Grammont, On the numerical analysis of generalized Sylvester equations. Part II: Theoretical sensitivity and numerical experiments, *Numer. Funct. Anal. and Optimiz.*, **15**, No-s: 7 and 8 (1994), 729-754.

[10] J.P. Brans, B. Mareschal, PROMCALC and Gaia: A new decision support system for multicriteria decision aid, *Decision Support Systems*, **12** (1994), 297-310.

[11] J.P. Brans, Ph. Vincke, B. Mareschal, How to select and how to rank projects: The Promethee method, *European journal of Operational Research*, **24** (1986), 228-238.

[12] D. Duvivier, V. Dhaevers, A. Artiba, Simulation-based performance measurement and analysis: an industrial application, *Internat. Journal of Computer Integrated Manufacturing (IJCIM)*, **18**, No. 5 (2005), 402-407.

[13] D. Duvivier, O. Roux, V. Dhaevers, A. Lazraq, Comparison of scheduling strategies via a statistical analysis in an industrial context, *International Conference on Service Systems and Service Management IEEE ICSSSM* (2006).

[14] D. Duvivier, O. Roux, V. Dhaevers, N. Meskens, A. Artiba, Multicriteria optimisation and simulation: An industrial application, *Annals of Operations Research*, **156**, No. 1 (2007), 45-60.

[15] J-C. Pomerol, S. Barba-Romero, *Multicriterion Decision Making in Management: Principles and Practice*, Kluwer, New York (2000).

[16] O. Roux, D. Duvivier, V. Dhaevers, N. Meskens, A. Artiba, A multicriteria approach to rank scheduling strategies, In: *International Conference on Industrial Engineering and Systems Management 2005 (IESM'05)*, Marrakech, Morocco (2005), 673-682.