Annexe E

Quelques rappels sur les matrices

Dans cette annexe d'abord nous rappelons un certain nombres de définitions et de relations sur les matrices et le calcul matriciel. Ensuite, les propriétés d'un certain nombres de matrices que l'on trouve souvent en traitement du signal et des images sont étudiées. parmi ces matrices, une place particulière est donnée aux matrices de Toeplitz, de Hankel, circulantes, bloc-Toeplitz et bloc-circulantes.

E.1 Introduction

E.1.1 Définition

Soient $\mathcal{V}$ et $\mathcal{W}$ deux espaces vectoriels sur le même corps (le corps $\mathcal{R}$ des nombres réels ou le corps $\mathcal{C}$ des nombres complexes ou, d'une façon plus générale, le corps $\mathcal{K}$ des scalaires), munis de bases $\{\boldsymbol{e}_i, i = 1, \cdots, m\}$ et $\{\boldsymbol{f}_j, j = 1, \cdots, n\}$ respectivement. Relativement à ces bases, une application linéaire

$$\boldsymbol{A} : \mathcal{V} \longrightarrow \mathcal{W}$$

est représentée par la matrice $\boldsymbol{A}$ à m lignes et n colonnes :

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

les éléments a_{ij} de la matrice $\boldsymbol{A}$ étant définis de façon unique par les relations

$$\boldsymbol{A}\boldsymbol{e}_j = \sum_{j=1}^{N} a_{ij} \boldsymbol{f}_j, \quad i = 1, \cdots, m.$$

Autrement dit, le j-ème vecteur colonne

$$\boldsymbol{a}_{*j} = [\boldsymbol{A}]_{*j} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

de la matrice $\boldsymbol{A}$ représente le vecteur $\boldsymbol{A}\boldsymbol{e}_i$ dans la base $\{\boldsymbol{f}_j, j = 1, \cdots, n\}$. On note

$$\boldsymbol{a}_{i*} = [\boldsymbol{A}]_{i*} = (\, a_{i1} \quad a_{i2} \quad \cdots \quad a_{in}\,)^t$$

le i-ème vecteur ligne de la matrice $\boldsymbol{A}$.

- Une matrice $\boldsymbol{A}$ d'éléments a_{ij} est notée $\boldsymbol{A} = [a_{ij}]$. Une matrice à m lignes et n colonnes est appelée matrice de type $[m, n]$, et on note $\mathcal{A}_{m,n}(\mathcal{K})$ l'espace vectoriel sur le corps $\mathcal{K}$ formé par les matrices de type $[m, n]$ à éléments dans $\mathcal{K}$. Un vecteur colonne est donc une matrice de type $[m, 1]$ et un vecteur ligne une matrice de type $[1, n]$.

- étant donnée une matrice $\boldsymbol{A} \in \mathcal{A}_{m,n}(\mathcal{C})$, on note $\boldsymbol{A}^* \in \mathcal{A}_{n,m}(\mathcal{C})$ la matrice adjointe de la matrice $\boldsymbol{A}$, définie de façon unique par les relations

$$< \boldsymbol{Au}, \boldsymbol{v} >_m = < \boldsymbol{u}, \boldsymbol{A}^*\boldsymbol{v} >_n, \quad \forall\, \boldsymbol{u} \in \mathcal{C}^n, \boldsymbol{v} \in \mathcal{C}^m$$

qui entraînent $[\boldsymbol{A}^*]_{ij} = a^*_{ji}$. De la même façon, étant donnée une matrice $\boldsymbol{A} \in \mathcal{A}_{m,n}(\mathcal{R})$, on note $\boldsymbol{A}^t \in \boldsymbol{A}_{n,m}(\mathcal{R})$ la matrice transposée de la matrice $\boldsymbol{A}$, définie de façon unique par les relations

$$< \boldsymbol{Au}, \boldsymbol{v} >_m = < \boldsymbol{u}, \boldsymbol{A}^*\boldsymbol{v} >_n \forall\, \boldsymbol{u} \in \mathcal{R}^n, \boldsymbol{v} \in \mathcal{R}^m$$

qui entraînent $[\boldsymbol{A}^t]_{ij} = a_{ji}$.

E.1.2 Opérations élémentaires

- **Addition :**

$$\boldsymbol{C} = \boldsymbol{A} + \boldsymbol{B}, \quad [\boldsymbol{C}]_{ij} = [\boldsymbol{A}]_{ij} + [\boldsymbol{B}]_{ij}$$

$$[\boldsymbol{A} + \boldsymbol{B} + \boldsymbol{C}] = [\boldsymbol{C} + \boldsymbol{B} + \boldsymbol{A}] = [\boldsymbol{B} + \boldsymbol{C} + \boldsymbol{A}]$$

- **Multiplication par un scalaire :**

$$\boldsymbol{C} = \lambda \boldsymbol{A}, \quad [\boldsymbol{C}]_{ij} = \lambda [\boldsymbol{A}]_{ij}$$

$$\lambda[\boldsymbol{A} + \boldsymbol{B}] = \lambda \boldsymbol{A} + \lambda \boldsymbol{B}$$

- **Multiplication de deux matrices :**

  À la composition des applications linéaires correspond la multiplication des matrices : si $\boldsymbol{A} = [a_{ij}]$ est une matrice de type $[I, J]$ et $B = [b_{jk}]$ est une matrice de type $[J, K]$, leur produit $\boldsymbol{AB}$ est la matrice $\boldsymbol{C} = [c_{ik}]$ de type $[I, K]$ définie par

$$\boldsymbol{C} = \boldsymbol{AB} \longrightarrow c_{ik} = \sum_{j=1}^{J} a_{ij} b_{jk}$$

On rappelle que l'on a :

$$[\boldsymbol{A}+\boldsymbol{B}]^t = \boldsymbol{A}^t + \boldsymbol{B}^t, \quad [\boldsymbol{A}+\boldsymbol{B}]^* = \boldsymbol{A}^* + \boldsymbol{B}^*$$

$$[\boldsymbol{AB}]^t = \boldsymbol{B}^t\boldsymbol{A}^t, \quad [\boldsymbol{AB}]^* = \boldsymbol{B}^*\boldsymbol{A}^*$$

$$\lambda[\boldsymbol{AB}] = [\lambda\boldsymbol{A}]\boldsymbol{B} = \boldsymbol{A}[\lambda\boldsymbol{B}]$$

$$\boldsymbol{A}[\boldsymbol{B}+\boldsymbol{C}] = \boldsymbol{AB} + \boldsymbol{AC}$$

$$[\boldsymbol{A}+\boldsymbol{B}]\boldsymbol{C} = \boldsymbol{AC} + \boldsymbol{BC}$$

E.1.3 Matrices élémentaires

- Une matrice de type $[n, n]$ est dite matrice carrée d'ordre n. Si $\boldsymbol{A}$ est une matrice carrée, les éléments a_{ii} sont appelés *éléments diagonaux*, et les éléments $a_{ij}, i \neq j$, sont appelés *éléments hors-diagonaux*.
- Une matrice de type $[n, n]$ est dite *triangulaire supérieure* si $a_{ij} = 0$ pour $i > j$, et *triangulaire inférieure* si $a_{ij} = 0$ pour $i < j$.
- La *matrice unité* est la matrice $\boldsymbol{I} = [d_{ij}]$.
- Une matrice est *diagonale* si $a_{ij} = 0$, pour $i \neq j$ on la note

$$\boldsymbol{A} = \mathrm{diag}\,[a_{ii}] = \mathrm{diag}\,[a_{11}, a_{22}, \cdots, a_{nn}]$$

- Une matrice $\boldsymbol{A}$ est :

symétrique	si	$\boldsymbol{A}$ est réelle et $\boldsymbol{A} = \boldsymbol{A}^t$
hermitienne	si	$\boldsymbol{A} = \boldsymbol{A}^*$
orthogonale	si	$\boldsymbol{A}$ est réelle et $\boldsymbol{AA}^t = \boldsymbol{A}^t\boldsymbol{A} = \boldsymbol{I}$
unitaire	si	$\boldsymbol{AA}^* = \boldsymbol{A}^*\boldsymbol{A} = \boldsymbol{I}$
normale	si	$\boldsymbol{AA}^* = \boldsymbol{A}^*\boldsymbol{A}$

E.1.4 Rang d'une matrice

- Le *rang* d'une matrice $\boldsymbol{A}$ de type $[M, N]$ est la dimension des sous espaces ligne et colonne engendrés respectivement par les vecteurs lignes $\{\boldsymbol{a}_{i*}, i = 1, \cdots, M\}$ et les vecteurs colonnes $\{\boldsymbol{a}_{*j}, j = 1, \cdots, N\}$ de $\boldsymbol{A}$.
- Une matrice est dite de *rang plein* lorsque $\mathrm{rang}\,[\boldsymbol{A}] = \min(M, N)$.
- $\mathrm{rang}\,[\boldsymbol{A}]$ est l'ordre de sa plus grande sous-matrice non singulière.

$$\mathrm{rang}\,[\boldsymbol{A}] = \mathrm{rang}\left[\boldsymbol{A}^t\right]$$

- Deux matrices sont dites équivalentes si elle ont le même rang.
- Si $\boldsymbol{B}$ est une matrice de rang plein régulière, alors les matrices $\boldsymbol{BA}$ ou $\boldsymbol{AB}$ ont le même rang que la matrice $\boldsymbol{A}$:

$$\mathrm{rang}\,[\boldsymbol{AB}] = \mathrm{rang}\,[\boldsymbol{BA}] = \mathrm{rang}\,[\boldsymbol{A}]$$

E.1.5 Matrice inverse

- Une matrice $\boldsymbol{A}$ est inversible s'il existe une matrice, notée $\boldsymbol{A}^{-1}$ et appelée matrice inverse de la matrice $\boldsymbol{A}$, telle que $\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}$. Dans le cas contraire, on dit que la matrice est singulière.
- Si $\boldsymbol{A}$ et B sont des matrices inversibles on a :

$$[\boldsymbol{A}^t]^{-1} = [\boldsymbol{A}^{-1}]^t, \quad [\boldsymbol{A}^*]^{-1} = [\boldsymbol{A}^{-1}]^* \quad \text{et} \quad [\boldsymbol{A}\boldsymbol{B}]^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$$

- Une CNS pour qu'une matrice carrée de dimensions $[N, N]$ soit inversible (ou régulière) est qu'elle soit de rang plein, c'est à dire $\text{rang}\,[\boldsymbol{A}] = N$.

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}, \qquad \boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{y}$$

$$(\boldsymbol{A}\boldsymbol{B}\boldsymbol{C})^{-1} = \boldsymbol{C}^{-1}\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$$

E.1.6 Déterminant d'une matrice

$$\det[\boldsymbol{A}] = |\boldsymbol{A}| = \sum_{j=1}^{N} [\boldsymbol{A}]_{ij} \boldsymbol{A}^{\ddagger}_{ij}$$

$$\det[\boldsymbol{A}] = \det\left[\boldsymbol{A}^t\right]$$

$$\det[\boldsymbol{A}\boldsymbol{B}\boldsymbol{C}] = \det[\boldsymbol{A}]\,\det[\boldsymbol{B}]\,\det[\boldsymbol{C}]$$

E.1.7 Matrice singulière

$\boldsymbol{A}$ est singulière si $\det[\boldsymbol{A}] = 0$.

$$\boldsymbol{A}\boldsymbol{x} = 0, \quad \boldsymbol{A}^t\boldsymbol{x} = \boldsymbol{0} \quad \forall \quad \boldsymbol{x} \neq \boldsymbol{0}.$$

E.1.8 Trace d'une matrice

La trace d'une matrice est définie par

$$\text{tr}\,[\boldsymbol{A}] = \sum_{i=1}^{N} a_{ii}$$

On a les relations suivantes:

$$\text{tr}\,[\boldsymbol{A}] = \sum_{i=1}^{N} a_{ii} = \sum_{i=1}^{N} \lambda_i$$

$$\text{tr}\,[\boldsymbol{A}\boldsymbol{B}\boldsymbol{C}] = \text{tr}\,[\boldsymbol{C}\boldsymbol{A}\boldsymbol{B}] = \text{tr}\,[\boldsymbol{B}\boldsymbol{C}\boldsymbol{A}]$$

$$\text{tr}\,[\boldsymbol{A} + \boldsymbol{B} + \boldsymbol{C}] = \text{tr}\,[\boldsymbol{A}] + \text{tr}\,[\boldsymbol{B}] + \text{tr}\,[\boldsymbol{C}]$$

La trace est invariante dans un changement de base.

E.1.9 Vecteurs propres et valeurs propre d'une matrice

- Un vecteur $\boldsymbol{u}$ est dit un *vecteur propre* de la matrice $\boldsymbol{A}$ si

$$\boldsymbol{A}\boldsymbol{u} = \lambda \boldsymbol{u};$$

  λ est alors une *valeurs propre* associée au vecteur propre $\boldsymbol{u}$.

- Pour une matrice carrée on définit :
  polynôme caractéristique: $f(\lambda) = \det[\boldsymbol{A} - \lambda \boldsymbol{I}]$
  Matrice caractéristique: $\boldsymbol{A} - \lambda \boldsymbol{I}$
  Équation caractéristique: $f(\lambda) = \det[\boldsymbol{A} - \lambda \boldsymbol{I} = 0]$

- Les *valeurs propres* $\{\lambda_i = \lambda_i(\boldsymbol{A}), i = 1, \cdots, n\}$ d'une matrice $\boldsymbol{A}$ d'ordre n sont les n racines, réelles ou complexes, distinctes ou confondues, du polynôme caractéristique

$$f(\lambda) = \det[\boldsymbol{A} - \lambda \boldsymbol{I}] = (-\lambda)^n + \sum_{i=1}^{N} a_{ij} (-\lambda)^{n-1} + \cdots + \det[\boldsymbol{A}] = 0$$

- Les v.p. d'une matrice ont une interprétation géométrique simple: Soit $\mathcal{E}^n$ une espace linéaire complexe à n dimensions et $\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n\}$ une base dans cet espace. A la matrice $\boldsymbol{A}$ et à cette base on peut associer un opérateur linéaire sous la forme de

$$\boldsymbol{A}\,\boldsymbol{e}_j = \sum_{i=1}^{N} a_{ji}\,\boldsymbol{e}_i, \quad j = 1, \cdots, n$$

  Ainsi les λ_i sont les v.p. de l'opérateur $\boldsymbol{A}$, c'est à dire qu'il existe un vecteur

$$\boldsymbol{u} = u_1 e_1 + \cdots + u_i e_i + \cdots + u_n e_n \neq \boldsymbol{0},$$

  appartenant à l'espace $\mathcal{E}^n$ de façon à avoir $\boldsymbol{A}\boldsymbol{u} = \lambda \boldsymbol{u}$. Ce vecteur est appelé le vecteur propre de l'opérateur $\boldsymbol{A}$ correspondant à la v.p. λ.

- On montre facilement que

$$\det[\boldsymbol{A}] = \prod_{i=1}^{n} \lambda_i$$

  Une matrice est régulière ssi $\lambda_i \neq 0, i = 1, \cdots, n$.

- Le *spectre* de la matrice $\boldsymbol{A}$ est le sous-ensemble

$$\mathrm{sp}[\boldsymbol{A}] = \cap_{j=1}^{n} \{\lambda_i(\boldsymbol{A})\}$$

  du plan complexe.

- Le *rayon spectral* d'une matrice $\boldsymbol{A}$ est le nombre $\rho(\boldsymbol{A}) \geq 0$ défini par

$$\rho(\boldsymbol{A}) = \max_i \{|\lambda_i(\boldsymbol{A})|, \quad i = 1, \cdots, n\}$$

- A tout valeur propre λ d'une matrice $\boldsymbol{A}$ est associé (au moins) un vecteur propre $\boldsymbol{v} \neq 0$ tel que : $\boldsymbol{A}\boldsymbol{v} = \lambda \boldsymbol{v}$.

- Si $\lambda \in \mathrm{sp}[\boldsymbol{A}]$, le sous-espace vectoriel $\{\boldsymbol{v} \in \mathcal{V}; \boldsymbol{A}\boldsymbol{v} = \lambda \boldsymbol{v}\}$ est appelé sous-espace propre correspondant à la valeur propre λ.

E.1.10 Valeurs singulières d'une matrice

Les valeurs singulières de la matrice $\boldsymbol{A}$ de dimensions $[M,N]$ sont données par

$$\mu_i^2 \boldsymbol{A}\boldsymbol{A}^* \boldsymbol{u}_i = \boldsymbol{u}_i, \quad i = 1, \cdots, r$$

$$\mu_i^2 \boldsymbol{A}^* \boldsymbol{A} \boldsymbol{v}_i = \boldsymbol{v}_i, \quad i = 1, \cdots, r$$

où $r = \text{rang}[\boldsymbol{A}]$. Si on définit $\lambda_i = \frac{1}{(\mu_i)^2}, \quad i = 1, \cdots, r$ en supposant $\lambda_1 \geq \lambda_2 \geq \cdots \geq 0$, et si on les complète par des zéros, c'est à dire $\{\lambda_i = 0, \quad i = r+1, \cdots, \max(M,N)\}$, alors on peut associer deux systèmes de v.p. $\{\boldsymbol{u}_i, \quad i = 1, \cdots, M\}$ et $\{\boldsymbol{v}_j, \quad j = 1, \cdots, N\}$ aux $\{\lambda_i\}$, tels que

$$\boldsymbol{A}\boldsymbol{u}_i = \lambda_i \boldsymbol{v}_i$$

$$\boldsymbol{A}\boldsymbol{v}_i = \lambda_i \boldsymbol{u}_j$$

$$\boldsymbol{A}\boldsymbol{A}^* \boldsymbol{u}_i = \lambda_i^2 \boldsymbol{u}_i, \quad i = 1, \cdots, M$$

$$\boldsymbol{A}^* \boldsymbol{A} \boldsymbol{v}_i = \lambda_i^2 \boldsymbol{v}_i, \quad j = 1, \cdots, N.$$

$\lambda_i^2, \boldsymbol{u}_i$ et $\lambda_i^2, \boldsymbol{v}_i$ constituent les ystèmes caractéristiques de $\boldsymbol{A}\boldsymbol{A}^*$ et $\boldsymbol{A}^*\boldsymbol{A}$, respectivement. La matrice $\boldsymbol{A}$ peut se factoriser en le produit

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{V}^t$$

où $\boldsymbol{U}$ est la matrice dont les colonnes sont les v.p. $\boldsymbol{u}_j$, $\boldsymbol{V}$ est la matrice dont les colonnes sont les v.p. $\boldsymbol{v}_i$ et $\boldsymbol{\Lambda}$ est la matrice diagonale formée de v.p. λ_i.

Le rang r de la matrice normale $\boldsymbol{A}^*\boldsymbol{A}$ est inférieur ou égale à $\inf(M,N)$. On complète les matrices $\boldsymbol{U}$ et $\boldsymbol{V}$ par les $(M-r)$ v.p. $\boldsymbol{u}_i$ et les $(N-r)$ v.p. $\boldsymbol{v}_i$ qui correspondent aux v.p. nulles, de façon à obtenir des bases complètes respectivement de Y et de X.

E.1.11 Matrices et les opérations linéaires

Si commençant par une base $\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n\}$, on introduit dans un espace euclidienne $\mathcal{E}^n$ un produit scalaire sous forme de

$$< \boldsymbol{x}, \boldsymbol{y} >= x_1 y_1^* + \cdots + x_i y_i^* + \cdots + x_n y_n^*$$

alors une matrice hermitienne $\boldsymbol{A}$: $([\boldsymbol{A}]_{ji} = [\boldsymbol{A}]_{ij}^*)$ induit sur cette base un opérateur hermitienne, c'est à dire

$$< \boldsymbol{A}\boldsymbol{x}, \boldsymbol{y} >=< \boldsymbol{x}, \boldsymbol{A}\boldsymbol{y} > \quad \forall \boldsymbol{x}, \boldsymbol{y} \in \mathcal{E}^n$$

– Le produit scalaire a les propriétés suivantes:

$$\forall \boldsymbol{x}, \boldsymbol{x}_1, \boldsymbol{x}_1, \boldsymbol{x}, \boldsymbol{y} \quad \text{et} \quad \boldsymbol{y}_2 \in \mathcal{E}^n$$

$$< \boldsymbol{x}, \boldsymbol{x} > > 0 \quad \forall \boldsymbol{x} \neq 0$$

$$< (\boldsymbol{x}_1 + \boldsymbol{x}_2), \boldsymbol{x} >=< \boldsymbol{x}_1, \boldsymbol{y} > + < \boldsymbol{x}_2, \boldsymbol{y} >$$

$$< \alpha \boldsymbol{x}, \boldsymbol{y} >= \alpha < \boldsymbol{x}, \boldsymbol{y} >$$

$$< \boldsymbol{x}, \boldsymbol{y} >=< \boldsymbol{y}, \boldsymbol{x} >^*$$

– Si $\boldsymbol{A}$ est une matrice hermitienne on a

$$< \boldsymbol{A}\boldsymbol{x}, \boldsymbol{x} >= \sum_i a_{ij} x_i x_j^*$$

Un telle forme est entièrement déterminée par une matrice hermitienne $\boldsymbol{A}$. L'ordre N et le rang de la matrice $\boldsymbol{A}$ sont appelés respectivement l'ordre et le rang de la forme $< \boldsymbol{A}\boldsymbol{x}, \boldsymbol{x} >$.

– Une tache importante de la théorie des formes hermitiennes est la réduction de la forme hermitienne à la somme de carrées suivantes

$$< \boldsymbol{A}\boldsymbol{x}, \boldsymbol{x} >= \sum_k \lambda_k |x_k|^2$$

où $x_k = c_{1k}x_1 + c_{2k}x_2 + \cdots$

Si r est le rang de $\boldsymbol{A}$ et x_i sont linéairement indépendantes, alors seul r coefficients λ_k dans cette équation son non-nulle.

– Soient $\mathcal{V}$ un espace vectoriel de dimension finie n, et $\mathcal{A} : \mathcal{V} \longrightarrow \mathcal{V}$ une application linéaire, représentée par une matrice carrée $\boldsymbol{A}$ relativement à une base $\{\boldsymbol{e}_i, i = 1, \cdots, m\}$. Cette même application, relativement à une autre base $\{\boldsymbol{f}_j, j = 1, \cdots, n\}$, est représentée par la matrice

$$\boldsymbol{B} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P},$$

où $\boldsymbol{P}$ est la matrice inversible dont le j-ème vecteur colonne est formé des composantes du vecteur $\boldsymbol{f}_j$ dans la base $\{\boldsymbol{e}_i, i = 1, \cdots, m\}$. La matrice $\boldsymbol{P}$ est appelée *matrice de passage de la base* $\{\boldsymbol{e}_j, j = 1, \cdots, n\}$ dans la base $\{\boldsymbol{f}_j, j = 1, \cdots, n\}$. Les deux matrices $\boldsymbol{A}$ et $\boldsymbol{B}$ sont dites *semblables* s'il existe une matrice inversible $\boldsymbol{P}$ telle que $\boldsymbol{B} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$.

– Une même application linéaire $\boldsymbol{A}$ étant représentée par différentes matrices selon la base choisie, le problème de la réduction des matrices consiste alors de trouver une base vis-à-vis de laquelle la matrice représentant l'application soit aussi simple que possible. Le cas le plus favorable est celui où il existe une matrice inversible $\boldsymbol{P}$ telle que la matrice $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ soit diagonale. Dans ce cas on dit que la matrice $\boldsymbol{A}$ est *diagonalisable*. Les éléments diagonaux de la matrice $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ sont alors les valeurs propres $\lambda_1, \lambda_2, \cdots, \lambda_n$ de la matrice $\boldsymbol{A}$, et le j-ème vecteur colonne $\boldsymbol{p}_j$ de la matrice $\boldsymbol{P}$ est formé des composantes d'un vecteur propre correspondant à λ_j, et on a :

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \operatorname{diag}[\lambda_i] \longrightarrow \boldsymbol{A}\boldsymbol{p}_j = \lambda_j \boldsymbol{p}_j, \quad j = 1, \cdots, n$$

Autrement dit, une matrice est diagonalisable ssi il existe une base de vecteur propres.

- Étant donnée une matrice carrée $\boldsymbol{A}$, il existe une matrice unitaire $\boldsymbol{U}$ telle que la matrice $\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}$ soit triangulaire.
- Étant donnée une matrice normale $\boldsymbol{A}$, il existe une matrice unitaire $\boldsymbol{U}$ telle que la matrice $\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}$ soit diagonale.
- Étant donnée une matrice symétrique $\boldsymbol{A}$, il existe une matrice orthogonale $\boldsymbol{O}$ telle que la matrice $\boldsymbol{O}^{-1}\boldsymbol{A}\boldsymbol{O}$ soit diagonale.
- On appelle *valeurs singulières* d'une matrice $\boldsymbol{A}$ carrée les racines carrées positives des valeurs propres de la matrice hermitienne $\boldsymbol{A}^*\boldsymbol{A}$ (ou $\boldsymbol{A}^t\boldsymbol{A}$ si la matrice $\boldsymbol{A}$ est réelle). Les valeurs singulières d'une matrice sont toujours ≥ 0. Elle sont toutes > 0 ssi la matrice $\boldsymbol{A}$ est inversible.
- Deux matrices $\boldsymbol{A}$ et $\boldsymbol{B}$ de type $[m, n]$ sont dites *équivalentes* s'il existe une matrice inversible $\boldsymbol{Q}$ d'ordre m et une matrice inversible $\boldsymbol{P}$ d'ordre n telle que $\boldsymbol{B} = \boldsymbol{Q}\boldsymbol{A}\boldsymbol{P}$.
- Pour une matrice réelle carrée $\boldsymbol{A}$, il existe deux matrices orthogonales $\boldsymbol{U}$ et $\boldsymbol{V}$ telle que
$$\boldsymbol{U}^t\boldsymbol{A}\boldsymbol{U} = \operatorname{diag}[\mu_i].$$
- Pour une matrice complexe carrée $\boldsymbol{A}$, il existe deux matrices unitaires $\boldsymbol{U}$ et $\boldsymbol{V}$ telle que
$$\boldsymbol{U}^*\boldsymbol{A}\boldsymbol{U} = \operatorname{diag}[\mu_i].$$
- Dans les deux cas, les nombres $\mu_i \geq 0$ sont les valeurs singulières de la matrice $\boldsymbol{A}$.
- Toutes les valeurs propres d'une matrice hermitienne sont réelles. Toute matrice hermitienne est diagonalisable et la matrice de passage sont unitaires.
- Soit $\boldsymbol{A}$ une matrice carrée représentant une application linéaire d'un espace $\mathcal{V}$ sur le corps $\mathcal{C}$, muni de son produit *canonique*.
- Le *quotient de Rayleigh* de la matrice $\boldsymbol{A}$ est l'application :
$$R_A : \{\boldsymbol{V} - \{0\}\} \longrightarrow \mathcal{C}$$
définie par :
$$R_A(\boldsymbol{v}) = \frac{<\boldsymbol{v}, \boldsymbol{A}\boldsymbol{v}>}{<\boldsymbol{v}, \boldsymbol{v}>} = \frac{\boldsymbol{v}^*\boldsymbol{A}\boldsymbol{v}}{\boldsymbol{v}^*\boldsymbol{v}}, \quad \boldsymbol{v} \neq 0$$

E.1.12 Décomposition tronquée en valeurs singulière

Si les λ_i décroissent rapidement alors on peut approcher la matrice $\boldsymbol{A}$ par une décomposition tronquée en valeurs singulières ($k < r = \text{rang}[\boldsymbol{A}]$) :

$$\boldsymbol{A} = \sum_{i=1}^{k} \lambda_i \boldsymbol{u}_i \boldsymbol{v}_i^* + \boldsymbol{E}_k$$

où $\boldsymbol{E}_k$ est la matrice d'erreur de l'approximation. Si on définit J_k comme l'énergie de l'erreur par

$$J_k = \sum_{i=1}^{N} \sum_{j=1}^{N} |[\boldsymbol{E}_k]_{ij}|^2 = \sum_{i=1}^{N} \sum_{j=1}^{N} [\boldsymbol{E}_k]_{ij} [\boldsymbol{E}_k]_{ij}^*$$

alors on montre facilement que l'on a

$$J_k = \sum_{i=k+1}^{r} \lambda_i$$

Si on définit α_k par

$$\alpha_k = \frac{J_k}{J_0} = 1 - \frac{\sum_{i=1}^{k} \lambda_i}{\sum_{i=1}^{r} \lambda_i}$$

alors $\quad 0 \leq \alpha_k \leq 1$, $\alpha_0 = 1$, et $\alpha_r = 0$.

E.1.13 Inverse généralisée d'une matrice

Considérons l'équation

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$$

où $\boldsymbol{A}$ est une matrice de dimension $[m, n]$.

- **Définition de Moore :**
  Une matrice $\boldsymbol{G}$ de dimension $[n, m]$ est dite inverse généralisée de la matrice $\boldsymbol{A}$ si

  $$\boldsymbol{A}\boldsymbol{G} = \boldsymbol{P}_A \quad \text{et} \quad \boldsymbol{G}\boldsymbol{A} = \boldsymbol{P}_G$$

  où $\boldsymbol{P}_A$ est l'opérateur (matrice) de projection des vecteurs de l'espace $\mathcal{E}^m$ sur $\mu(\boldsymbol{A})$ et $\boldsymbol{P}_G$ est l'opérateur (matrice) de projection des vecteurs de l'espace $\mathcal{E}^n$ sur $\mu(\boldsymbol{G})$.

- **Définition de Penrose :**
  Une matrice $\boldsymbol{G}$ de dimension $[n, m]$ est dite inverse généralisée de la matrice $\boldsymbol{A}$ si

  $$\boldsymbol{A}\boldsymbol{G}\boldsymbol{A} = \boldsymbol{A}, \quad (\boldsymbol{A}\boldsymbol{G})^* = \boldsymbol{A}\boldsymbol{G}, \quad \boldsymbol{G}\boldsymbol{A}\boldsymbol{G} = \boldsymbol{G} \quad \text{et} \quad (\boldsymbol{G}\boldsymbol{A})^* = \boldsymbol{G}\boldsymbol{A}$$

- Ces définitions sont identiques si les normes associées aux espaces $\mathcal{E}^m$ et $\mathcal{E}^n$ sont du type euclidien: $\|\boldsymbol{x}\| = <\boldsymbol{x}^*, \boldsymbol{x}>^{1/2}$.

E.2 Quelques matrices célèbres

Matrices identité :

$$[\boldsymbol{I}]_{ij} = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & \text{ailleurs} \end{cases}, \quad j = 1, \cdots, N$$

$$\boldsymbol{I}^t \boldsymbol{I} = \boldsymbol{I}\boldsymbol{I}^t = \boldsymbol{I}, \quad \boldsymbol{I}\boldsymbol{A} = \boldsymbol{A}\boldsymbol{I} = \boldsymbol{A}$$

Matrices diagonales :

$$\boldsymbol{D} = \text{diag}\,[d_1, \cdots, d_N] \longrightarrow [\boldsymbol{D}]_{ij} = d_i\,\delta_{ij}$$

$$\boldsymbol{D}^{-1} = \text{diag}\left[\frac{1}{d_1}, \cdots, \frac{1}{d_N}\right]$$

Matrices d'inversion d'ordre

$$\boldsymbol{J} = \begin{pmatrix} . & . & . & . & 1 \\ . & 0 & . & 1 & . \\ . & . & 1 & . & . \\ . & 1 & . & 0 & . \\ 1 & . & . & . & \end{pmatrix} \quad [\boldsymbol{J}]_{ij} = \begin{cases} 1 & i = N - j + 1, \quad j = 1, \cdots, N \\ 0 & \text{ailleurs} \end{cases}$$

$$[\boldsymbol{J}]^{2k} = \boldsymbol{I}, \quad [\boldsymbol{J}]^{2k+1} = \boldsymbol{J}, \quad \boldsymbol{J}^{-1} = \boldsymbol{J}$$

- $\boldsymbol{AJ}$ inverse l'ordre des éléments des lignes de la matrice $\boldsymbol{A}$.
- $\boldsymbol{JA}$ inverse l'ordre des éléments des colonnes de la matrice $\boldsymbol{A}$.
- $\boldsymbol{JAJ}$ inverse l'ordre des éléments des lignes et des colonnes de la matrice $\boldsymbol{A}$.
- $\boldsymbol{Jh}$ inverse l'ordre des éléments du vecteur $\boldsymbol{h}$.
- Si $\boldsymbol{h}$ est un vecteur symétrique on a $\boldsymbol{Jh} = \boldsymbol{h}$.
- Si $\boldsymbol{h}$ est un vecteur anti-symétrique on a $\boldsymbol{Jh} = -\boldsymbol{h}$.

Matrices symétriques :

$$[\boldsymbol{S}]_{ij} = [\boldsymbol{S}]_{ji}, \quad i, j = 1, \cdots, N, \quad \boldsymbol{S}^t = \boldsymbol{S}$$

- Si $\boldsymbol{S}$ est une matrice symétrique avec des éléments réels, alors ses v.p. sont réelles et il existe au moins une matrice $\boldsymbol{R}$ permettant d'avoir: $\boldsymbol{R}^t\boldsymbol{S}\boldsymbol{R} = \boldsymbol{D}$, $\boldsymbol{S} = \boldsymbol{R}\boldsymbol{D}\boldsymbol{R}^t$, où $\boldsymbol{D}$ est une matrice diagonale.
- Étant donnée une matrice symétrique $\boldsymbol{A}$, il existe une matrice orthogonale $\boldsymbol{O}$ telle que la matrice $\boldsymbol{O}^{-1}\boldsymbol{A}\boldsymbol{O}$ soit diagonale.

Matrices anti-symétriques :

$$[\boldsymbol{R}]_{ij} = -[\boldsymbol{R}]_{ji}, \quad i,j = 1,\cdots,N, \quad i \neq j \quad \boldsymbol{R}^t = -\boldsymbol{R}$$

Matrices symétriques par rapport à la deuxième diagonale :

$$[\boldsymbol{S}]_{ij} = [\boldsymbol{S}]_{(N-j+1)(N-i+1)}, \quad i,j = 1,\cdots,N, \quad \boldsymbol{JSJ} = \boldsymbol{S}^t, \quad \boldsymbol{JS}^t\boldsymbol{J} = \boldsymbol{S}$$

Matrices anti-symétriques par rapport à la deuxième diagonale :

$$[\boldsymbol{R}]_{ij} = -[\boldsymbol{R}]_{(N-i+1)(N-j+1)}, \quad i,j = 1,\cdots,N, \quad i \neq j, \quad \boldsymbol{JSJ} = -\boldsymbol{S}^t$$

Matrices hermitienne :

$$[\boldsymbol{H}]_{ij} = [\boldsymbol{H}^*]_{ji}, \quad i,j = 1,\cdots,N, \quad \boldsymbol{H}^* = \boldsymbol{H}^*$$

$[\boldsymbol{H}]_{ii}$ sont réels. Si $\boldsymbol{H}$ est une matrice hermitienne $\boldsymbol{H}^{-1}$ (à condition qu'elle existe) est aussi hermitienne, et on a $\boldsymbol{x}^t\boldsymbol{H}\boldsymbol{x} = \alpha\|\boldsymbol{x}\|^2$ est réel. $\boldsymbol{A}$ toute matrice hermitienne on peut toujours associer un système de vecteurs propres orthonormés complet. Les v.p. de $\boldsymbol{H}$ sont réelles et ses vecteurs propres sont orthogonaux.

Matrices anti-hermitiennes :

$$[\boldsymbol{K}]_{ij} = -[\boldsymbol{K}^*]_{ji}, \quad i,j = 1,\cdots,N, \quad \boldsymbol{K}^* = -\boldsymbol{K}^*$$

Si $\boldsymbol{K}$ est une matrice anti-hermitienne $\boldsymbol{K}^{-1}$ (à condition qu'elle existe) est aussi anti-hermitienne, et on a $\boldsymbol{x}^t\boldsymbol{K}\boldsymbol{x} = -j\alpha\|\boldsymbol{x}\|^2$ est imaginaire Les v.p. de $\boldsymbol{K}$ sont imaginaires et ses vecteurs propres sont orthogonaux. Une matrice quelconque $\boldsymbol{A}$ peut toujours être décomposée de la façon suivante :

$$\boldsymbol{A} = \frac{1}{2}[\boldsymbol{A} + \boldsymbol{A}^t] + \frac{1}{2}[\boldsymbol{A} - \boldsymbol{A}^t] = \frac{1}{2}[\boldsymbol{H} + \boldsymbol{K}].$$

Matrices idempotentes :

$$\boldsymbol{Q}^t\boldsymbol{Q} = \boldsymbol{Q}$$

Si $\boldsymbol{A}$ est une matrice idempotente et symétrique on a:

$$\boldsymbol{A}^2 = \boldsymbol{A}^t\boldsymbol{A} = \boldsymbol{A}$$

$$\boldsymbol{A}^2\boldsymbol{x}_i = \boldsymbol{A}\boldsymbol{x}_i = \lambda_i\boldsymbol{x}_i = \boldsymbol{A}(\lambda_i\boldsymbol{x}_i) = \lambda_i^2\boldsymbol{x}_i \longrightarrow \lambda_i = 0 \text{ ou } 1.$$

Matrices orthogonales :

$$\boldsymbol{O}^t\boldsymbol{O} = \boldsymbol{O}\boldsymbol{O}^t = \boldsymbol{O}\boldsymbol{O}^{-1} = \boldsymbol{O}^{-1}\boldsymbol{O} = \boldsymbol{I}$$

Deux matrices $\boldsymbol{A}$ et $\boldsymbol{B}$ sont *similaires* s'il existe une matrice $\boldsymbol{C}$ telle que

$$\boldsymbol{B} = \boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}$$

On a alors : $\boldsymbol{A} = \boldsymbol{C}^{-1}\boldsymbol{B}\boldsymbol{C}$. Deux matrices $\boldsymbol{A}$ et $\boldsymbol{B}$ sont *orthogonalement similaires* ou *congruentes* s'il existe une matrice orthogonale $\boldsymbol{O}$ telle que $\boldsymbol{B} = \boldsymbol{O}^{-1}\boldsymbol{A}\boldsymbol{O}$. On a alors $\boldsymbol{A} = \boldsymbol{O}^{-1}\boldsymbol{B}\boldsymbol{O}$.

Matrices de décalages :

$\boldsymbol{Z}$ est une matrices de décalage vers le bas (ou vers la droite) si :

$$\boldsymbol{Z} = \begin{pmatrix} 0 & . & . & . & 0 \\ 1 & 0 & . & . & 0 \\ 0 & 1 & 0 & . & 0 \\ . & 0 & 1 & 0 & 0 \\ . & . & 0 & 1 & 0 \end{pmatrix}, \quad [\boldsymbol{Z}]_{ij} = \begin{cases} 1 & i = j+1 \\ 0 & \text{ailleurs} \end{cases}, \quad j = 1, \cdots N-1$$

$\boldsymbol{Z}^t$ est une matrices de décalage vers le haut.

Si $\boldsymbol{Z}$ est une matrice de décalage vers le haut on a:

$$\begin{array}{ll} \boldsymbol{B} = \boldsymbol{Z}^{\alpha}\boldsymbol{A}\boldsymbol{Z}^{\beta} & [\boldsymbol{B}]_{mn} = [\boldsymbol{A}]_{m+\alpha n-\beta} \\ \boldsymbol{B} = \boldsymbol{Z}^{\alpha}\boldsymbol{A}\boldsymbol{Z}^{t\beta} & [\boldsymbol{B}]_{mn} = [\boldsymbol{A}]_{m+\alpha n+\beta} \\ \boldsymbol{B} = \boldsymbol{Z}^{t\alpha}\boldsymbol{A}\boldsymbol{Z}^{\beta} & [\boldsymbol{B}]_{mn} = [\boldsymbol{A}]_{m-\alpha n-\beta} \\ \boldsymbol{B} = \boldsymbol{Z}^{t\alpha}\boldsymbol{A}\boldsymbol{Z}^{t\beta} & [\boldsymbol{B}]_{mn} = [\boldsymbol{A}]_{m-\alpha n+\beta} \end{array}$$

Matrices de décalage circulaire :

$\boldsymbol{P}_0$ est une matrice de décalage circulaire vers le haut (ou vers la gauche) si :

$$\boldsymbol{P}_0 = \begin{pmatrix} 0 & 1 & 0 & . & 0 \\ 0 & 0 & 1 & 0 & . \\ 0 & . & 0 & 1 & 0 \\ . & . & . & 0 & 1 \\ 1 & . & . & . & 0 \end{pmatrix}, \quad [\boldsymbol{P}_0]_{ij} = \begin{cases} 1 & i = 1, \cdots, N-1 \\ 1 & j = 1, i = N \\ 0 & \text{ailleurs} \end{cases}$$

$\boldsymbol{P}_0$ est une matrice circulante. Elle peut donc être factorisée de façon suivante:

$$\boldsymbol{P}_0 = \boldsymbol{W} D \boldsymbol{W}^{-1}$$

où D est une matrice diagonale et les éléments des matrices $\boldsymbol{W}$ et $\boldsymbol{W}^{-1}$ sont données par:

$$[\boldsymbol{W}]_{pq} = \exp\left[\frac{2\pi}{N}(p-1)(q-1)\right] p, q = 1, \cdots, N$$

$$[\boldsymbol{W}^{-1}]_{pq} = \frac{1}{N} \exp\left[\frac{-2\pi}{N}(p-1)(q-1)\right], \quad p, q = 1, \cdots N$$

$\boldsymbol{P}_1 = \boldsymbol{P}_0^t$ est une matrices de décalage circulaire vers le bas (ou vers la droite).

$$\boldsymbol{P}_1^{-1} = \boldsymbol{P}_1^t = \boldsymbol{P}_0, \quad \boldsymbol{P}_0^{-1} = \boldsymbol{P}_0^t = \boldsymbol{P}_1$$

Matrices triangulaires supérieure :

$$\boldsymbol{U} = \begin{pmatrix} u_{11} & u_{12} & u_{13} & u_{14} & u_{15} \\ 0 & u_{22} & u_{23} & u_{24} & u_{25} \\ 0 & 0 & u_{33} & u_{34} & u_{35} \\ 0 & 0 & 0 & u_{44} & u_{45} \\ 0 & 0 & 0 & 0 & u_{55} \end{pmatrix}$$

$$\boldsymbol{U}_1 = \begin{pmatrix} 1 & u_{12} & u_{13} & u_{14} & u_{15} \\ 0 & 1 & u_{23} & u_{24} & u_{25} \\ 0 & 0 & 1 & u_{34} & u_{35} \\ 0 & 0 & 0 & 1 & u_{45} \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad \boldsymbol{U}_0 = \begin{pmatrix} 0 & u_{12} & u_{13} & u_{14} & u_{15} \\ 0 & 0 & u_{23} & u_{24} & u_{25} \\ 0 & 0 & 0 & u_{34} & u_{35} \\ 0 & 0 & 0 & 0 & u_{45} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

L'inverse d'une matrice du type $\boldsymbol{U}_1$ est aussi une matrice du type $\boldsymbol{U}_1$.

$$\det[\boldsymbol{U}] = \prod_{i=1}^{n} u_{ii}, \quad \det[\boldsymbol{U}_1] = 1, \quad \det[\boldsymbol{U}_0] = 0$$

Matrices triangulaires inférieure :

$$\boldsymbol{L} = \begin{pmatrix} l_{11} & 0 & 0 & 0 & 0 \\ l_{21} & l_{22} & 0 & 0 & 0 \\ l_{31} & l_{32} & l_{33} & 0 & 0 \\ l_{41} & l_{42} & l_{43} & l_{44} & 0 \\ l_{51} & l_{52} & l_{53} & l_{54} & l_{55} \end{pmatrix}$$

$$\boldsymbol{L}_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ l_{21} & 1 & 0 & 0 & 0 \\ l_{31} & l_{32} & 1 & 0 & 0 \\ l_{41} & l_{42} & l_{43} & 1 & 0 \\ l_{51} & l_{52} & l_{53} & l_{54} & 1 \end{pmatrix} \boldsymbol{L}_0 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ l_{21} & 0 & 0 & 0 & 0 \\ l_{31} & l_{32} & 0 & 0 & 0 \\ l_{41} & l_{42} & l_{43} & 0 & 0 \\ l_{51} & l_{52} & l_{53} & l_{54} & 0 \end{pmatrix}$$

L'inverse d'une matrice du type $\boldsymbol{L}_1$ est aussi une matrice du type $\boldsymbol{L}_1$.

$$\det[\boldsymbol{L}] = \prod_{i=1}^{n} l_{ii}, \quad \det[\boldsymbol{L}_1] = 1, \quad \det[\boldsymbol{L}_0] = 0$$

Matrices de Hessenberg supérieure :

$$\boldsymbol{H} = \begin{pmatrix} h_{11} & h_{12} & h_{13} & h_{14} & h_{15} \\ h_{12} & h_{22} & h_{23} & h_{24} & h_{25} \\ 0 & h_{23} & h_{33} & h_{34} & h_{35} \\ 0 & 0 & h_{34} & h_{44} & h_{45} \\ 0 & 0 & 0 & h_{45} & h_{55} \end{pmatrix}$$

Matrices de Hessenberg inférieure :

$$\boldsymbol{H} = \begin{pmatrix} h_{11} & h_{12} & 0 & 0 & 0 \\ h_{21} & h_{22} & h_{23} & 0 & 0 \\ h_{31} & h_{32} & h_{33} & h_{34} & 0 \\ h_{41} & h_{42} & h_{43} & h_{44} & h_{45} \\ h_{51} & h_{52} & h_{53} & h_{54} & h_{55} \end{pmatrix}$$

Matrices tridiagonale :

$$\boldsymbol{B} = \begin{pmatrix} b_{11} & b_{12} & 0 & 0 & 0 \\ b_{21} & b_{22} & b_{23} & 0 & 0 \\ 0 & b_{32} & b_{33} & b_{34} & 0 \\ 0 & 0 & b_{43} & b_{44} & b_{45} \\ 0 & 0 & 0 & b_{54} & b_{55} \end{pmatrix}$$

Matrice de transformation de Householder

Définition

Soit $\boldsymbol{v}$ un vecteur complexe de dimension $[N]$. Alors $\boldsymbol{v}^{*t}\boldsymbol{v}$ est un scalaire réel et $\boldsymbol{v}\boldsymbol{v}^{*t}$ est une matrice carrée de dimensions $[N, N]$. Soit $\boldsymbol{I}$ une matrice identité de mêmes dimensions. Alors la matrice

$$\boldsymbol{H_v} = \boldsymbol{I} - \frac{2}{\boldsymbol{v}^*\boldsymbol{v}}[\boldsymbol{v}\boldsymbol{v}^*]$$

est une matrice de réflexion ou matrice de Householder.

Propriétés :

- $\boldsymbol{H_v}$ est hermitienne et orthonormale.
- La transformation $\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x}$ avec $\boldsymbol{H}$ une matrice de Householder est appelée la transformation de Householder. Une telle transformation conserve l'énergie du vecteur original $\boldsymbol{x}$. (Ceci est vrai pour n'importe quelle transformation orthonormale.)
- Soit $\boldsymbol{e}_j$ un vecteur base ($[\boldsymbol{e}_j]i = \delta_{ij}$), alors pour un vecteur quelconque $\boldsymbol{u}$ il existe une matrice de Householder $\boldsymbol{H}$ telle que

  $$\boldsymbol{H}\boldsymbol{u} = \sigma \boldsymbol{e}_j$$

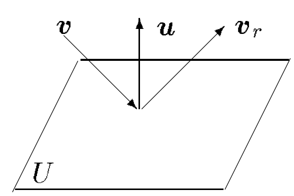
  ssi σ est choisi telle que

  $$\boldsymbol{u}^*\boldsymbol{u} = |\sigma^2| = \sum_{i=1}^{N} u_i u_i^* \quad \text{ou bien} \quad \sigma = \pm \frac{u_j}{|u_j|}\sqrt{\boldsymbol{u}^*\boldsymbol{u}}$$

  On a aussi $\boldsymbol{u}^*\boldsymbol{H} = \sigma^* \boldsymbol{e}_j^t$
- Soit $\boldsymbol{v}$ un vecteur, et $\boldsymbol{v}_r$ son image à travers un hyperplan U sous-espace orthogonale au vecteur $\boldsymbol{u}$ (symétrie plane). La matrice de transformation de Householder $\boldsymbol{H_u}$ est une matrice qui nous permet de faire correspondre $\boldsymbol{v}_r$ à $\boldsymbol{v}$ par

  $$\boldsymbol{v}_r = \boldsymbol{H_u}\, \boldsymbol{v}$$

  où $\boldsymbol{H_u} = \boldsymbol{I} - \beta \boldsymbol{u}\boldsymbol{u}^t$ avec $\beta = \frac{2}{(\boldsymbol{u}^t\boldsymbol{u})}$



Matrice de transformation de Householder hyperbolique

Soit $\boldsymbol{\Phi}$ une matrice diagonale avec des éléments diagonaux de $+1$ et 1. Soit $\boldsymbol{v}$ est un vecteur complexe. Alors

$$\boldsymbol{v}^*\Phi\boldsymbol{v} = \sum_{i=1}^{N} |v_i|^2[\Phi]_{ii}$$

est la norme hyperbolique du vecteur $\boldsymbol{v}$. Une matrice $\boldsymbol{Q}$ qui satisfait

$$\boldsymbol{Q\Phi Q}^* = \boldsymbol{\Phi}$$

est une matrice hypernormale. Une telle matrice préserve la norme hyperbolique des vecteurs, c'est à dire: si $\boldsymbol{y}^t = \boldsymbol{x}^t\boldsymbol{Q}$ alors $\boldsymbol{y}^t\boldsymbol{Q}\boldsymbol{y} = \boldsymbol{x}^t\boldsymbol{Q}\boldsymbol{x}$. Une matrice définie par

$$\boldsymbol{H} = \Phi - 2\frac{\boldsymbol{Q}\boldsymbol{Q}^t}{\boldsymbol{Q}^t\Phi\boldsymbol{Q}}$$

est appelée une matrice de transformation hyperbolique de Householder. $\boldsymbol{H}$ est hermitienne et hypernormale par rapport à la matrice $\boldsymbol{\Phi}$, c'est à dire $\boldsymbol{H\Phi H}^t = \boldsymbol{\Phi}$.

Propriétés :

- $\det[\boldsymbol{H}] = 1$; $\boldsymbol{H}$ est une matrice non-singulière.
- $\boldsymbol{H}$ n'est, en général, pas symétrique, mais elle a une symétrie hyperbolique, c'est à dire $\boldsymbol{H}^t\boldsymbol{A}\boldsymbol{H} = \boldsymbol{H}\boldsymbol{A}\boldsymbol{H}^t$.
- Les valeurs propres d'une matrice hypernormale forment des paires réciproquement conjuguées, c'est à dire que si λ est une valeur propre de cette matrice $\frac{1}{\lambda^*}$ l'est aussi. De plus si λ est multiple d'ordre n alors $\frac{1}{\lambda^*}$ est aussi de même ordre n.

Décomposition orthogonale par transformation de Householder

Soit $\boldsymbol{A}$ une matrice de dimensions $[N, M]$ et de rang $N \leq M$. Il existe une matrice orthogonale $\boldsymbol{H_u}$ de dimensions $[N, M]$ telle que $S = \boldsymbol{HuA}$ soit triangulaire supérieure. La matrice $\boldsymbol{A}$ peut être décomposée en le produit d'une matrice orthogonale $\boldsymbol{Hu}$ et d'une matrice triangulaire supérieure $\boldsymbol{S}$ de rang plein— $\boldsymbol{A} = \boldsymbol{HuS}$.

Matrices centro-symétriques

$$[\boldsymbol{G}]_{ij} = [\boldsymbol{G}]_{ji} = [\boldsymbol{G}]_{N+1-i,N+1-j}, \quad i,j = 1,\cdots,N$$

Exemple :

$$\boldsymbol{G} = \begin{pmatrix} g_{11} & g_{12} & g_{13} & g_{14} & g_{15} \\ g_{12} & g_{22} & g_{23} & g_{24} & g_{14} \\ g_{13} & g_{23} & g_{33} & g_{23} & g_{13} \\ g_{14} & g_{24} & g_{23} & g_{22} & g_{12} \\ g_{15} & g_{14} & g_{13} & g_{12} & g_{11} \end{pmatrix}$$

- Si $\boldsymbol{G}$ est une matrice centro-symétrique on a:

$$\boldsymbol{JG} = \boldsymbol{G}, \quad \boldsymbol{GJ} = \boldsymbol{G}, \quad \boldsymbol{JGJ} = \boldsymbol{G}, \quad \text{et} \quad \boldsymbol{JG}^{-1}\boldsymbol{J} = \boldsymbol{G}^{-1}$$

- Cette dernière relation montre que l'inverse d'une matrice centro-symétrique est aussi une matrice centro-symétrique.
- Si $\boldsymbol{G}$ est une matrice centro-symétrique de dimension $[N, N]$, alors elle peut être partitionnée de la manière suivante :
  - si N est pair ($N = 2r$) on a :

$$\boldsymbol{G} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{JB}^*\boldsymbol{J} & \boldsymbol{JA}^*\boldsymbol{J} \end{pmatrix}$$

  - si N est impair ($N = 2r + 1$) on a :

$$\boldsymbol{G} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{x} & \boldsymbol{B} \\ \boldsymbol{x}^* & \alpha & \boldsymbol{x}^*\boldsymbol{J} \\ \boldsymbol{JB}^*\boldsymbol{J} & \boldsymbol{Jx} & \boldsymbol{JA}^*\boldsymbol{J} \end{pmatrix}$$

  où $\boldsymbol{A}$ et $\boldsymbol{B}$ sont deux matrices de dimensions $[r, r]$, $\boldsymbol{J}$ est une matrice d'inversion d'ordre de dimensions $[r, r]$, $\boldsymbol{x}$ est un vecteur de dimension r et α est un scalaire.
- Si $\boldsymbol{G}$ est une matrice centro-symétrique de dimension $[N, N]$, alors
  - si N est pair on a

$$\boldsymbol{G} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{C}^t \\ \boldsymbol{C} & \boldsymbol{JAJ} \end{pmatrix}$$

  où $\boldsymbol{A}, \boldsymbol{C}$ et $\boldsymbol{J}$ sont des matrices de dimensions $[N/2, N/2]$ et

$$\boldsymbol{A}^t = \boldsymbol{A} \quad \text{et} \quad \boldsymbol{C}^t = \boldsymbol{JCJ}$$

  On montre aussi que $\boldsymbol{G}$ est orthogonalement similaire à et on montre que si $\boldsymbol{G}$ a des valeurs propres (v.p.) distinctes, elle a $[N/2]$ vecteurs propres symétriques ($\boldsymbol{h}_i, i = 1,\cdots,N/2$) et $[N/2)]$ vecteurs propres v.p. anti-symétriques ($\boldsymbol{g}_i, i = 1,\cdots,N/2$).

$$\begin{pmatrix} \boldsymbol{A} - \boldsymbol{JC} & 0 \\ 0 & \boldsymbol{A} + \boldsymbol{JC} \end{pmatrix}$$

Dans ce cas les v.p. symétriques $\boldsymbol{h}_i$ de $\boldsymbol{G}$ s'obtiennent par

$$\boldsymbol{h}_i = \frac{1}{\sqrt{2}}[\boldsymbol{u}_i, -[\boldsymbol{J}\boldsymbol{u}_i]^t]^t$$

De même les v.p. anti-symétriques $\boldsymbol{g}_i$ de $\boldsymbol{G}$ s'obtiennent par

$$\boldsymbol{g}_i = \frac{1}{\sqrt{2}}[\boldsymbol{v}_i, +[\boldsymbol{J}\boldsymbol{v}_i]^t]^t$$

où $\boldsymbol{u}_i$ et $\boldsymbol{v}_i$ sont respectivement les v.p. orthonormés des matrices $\boldsymbol{A} - \boldsymbol{J}\boldsymbol{C}$ et $\boldsymbol{A} + \boldsymbol{J}\boldsymbol{C}$.

– si N est impair on a

$$\boldsymbol{G} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{x} & \boldsymbol{C}^t \\ \boldsymbol{x}^t & q & \boldsymbol{x}^t\boldsymbol{J} \\ \boldsymbol{C} & \boldsymbol{J}\boldsymbol{x} & \boldsymbol{J}\boldsymbol{A}\boldsymbol{J} \end{pmatrix}$$

où $\boldsymbol{A}$, $\boldsymbol{C}$ et $\boldsymbol{J}$ sont des matrices de dimensions $[N/2-1, N/2-1]$ et

$$\boldsymbol{A}^t = \boldsymbol{A} \quad \text{et} \quad \boldsymbol{C}^t = \boldsymbol{J}\boldsymbol{C}\boldsymbol{J}$$

et on montre que si $\boldsymbol{G}$ a des valeurs propres distinctes elle a $[(N+1)/2)]$ v.p. symétriques $(\boldsymbol{h}_i, \quad i = 1, \cdots, (N+1)/2))$ et $[(N-1)/2)]$ v.p. anti-symétriques $(\boldsymbol{g}_i, \quad i = 1, \cdots, (N-1)/2))$

On montre aussi que $\boldsymbol{G}$ est orthogonalement similaire à

$$\begin{pmatrix} \boldsymbol{A} - \boldsymbol{J}\boldsymbol{C} & 0 & 0 \\ 0 & q & \sqrt{2}\boldsymbol{x}^t \\ 0 & \sqrt{2}\boldsymbol{x} & \boldsymbol{A} + \boldsymbol{J}\boldsymbol{C} \end{pmatrix}$$

Dans ce cas les v.p. symétriques $\boldsymbol{h}_i$ de $\boldsymbol{G}$ s'obtiennent par

$$\boldsymbol{h}_i = \frac{1}{\sqrt{2}}[\boldsymbol{u}_i^t, 0, -[\boldsymbol{J}\boldsymbol{u}_i]^t]^t$$

De même les v.p. anti-symétriques $\boldsymbol{g}_i$ de $\boldsymbol{G}$ s'obtiennent par

$$\boldsymbol{g}_i = \frac{1}{\sqrt{2}}[\boldsymbol{v}_i^t, \sqrt{2}\alpha_i, [\boldsymbol{J}\boldsymbol{v}_i]^t]^t$$

où $\boldsymbol{u}_i$ sont les v.p. orthonormés de la matrices $\boldsymbol{A} - \boldsymbol{J}\boldsymbol{C}$, et $[\boldsymbol{v}_i, \alpha_i]^t$ sont les v.p. orthonormés de la matrice

$$\begin{pmatrix} \boldsymbol{A} + \boldsymbol{J}\boldsymbol{C} & \sqrt{2}\boldsymbol{x} \\ \sqrt{2}\boldsymbol{x}^t & q \end{pmatrix}$$

Dans tous les cas ces v.p. sont uniques (à une constante multiplicative près) et mutuellement orthogonaux entre eux.

Si $\boldsymbol{a} = [a_0, a_1, \cdots, a_{(N-1)}]^t$ est un v.p. de $\boldsymbol{G}$, c'est à dire $\boldsymbol{G}\boldsymbol{a} = \lambda\boldsymbol{a}$, le polynôme

$$\boldsymbol{A}(z) = \sum_{k=0}^{N-1} a_k z^{-k}$$

est un polynôme d'ordre $(N-1)$, appelé polynôme propre de la matrice $\boldsymbol{G}$. On montre que si $z = z_i$ est une racine de $\boldsymbol{A}(z)$, alors $z = \frac{1}{z_i}$ en est une aussi.

Matrices d'innovation diagonales (MID)

(En Anglais : Diagonal Innovation matrices DIM)

Considérons les trois vecteurs $\boldsymbol{x}_{m-1}$ de dimension $[m-1]$, $\boldsymbol{z}_{m-1}$ de dimension $[m-1]$, et de $\boldsymbol{y}_m$ de dimension $[m]$. Une matrice MID est une matrice qui est générée par les trois vecteurs génératrices $\boldsymbol{x}_{m-1}$, $\boldsymbol{z}_{m-1}$, et $\boldsymbol{y}_m$ de façon suivante:

$$\boldsymbol{R} = \left(\begin{array}{ccccccccc} y_1 & \vdots & x_1 & \vdots & x_1 & \vdots & x_1 & \vdots & x_1 \\ \cdots & & & & & & & & \\ z_1 & & y_2 & \vdots & x_2 & \vdots & x_2 & \vdots & x_2 \\ \cdots & \cdots & \cdots & & & & & & \\ z_1 & & z_2 & & y_3 & \vdots & x_3 & \vdots & x_3 \\ \cdots & \cdots & \cdots & \cdots & \cdots & & & & \\ z_1 & & z_2 & & z_3 & & y_4 & \vdots & x_4 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & & & \\ z_1 & & z_2 & & z_3 & & z_4 & & y_5 \end{array}\right), \quad \left\{\begin{array}{l} \boldsymbol{x}_{m+1}^t = [\boldsymbol{x}_m^t, x_{m+1}] \\ \boldsymbol{z}_{m+1}^t = [\boldsymbol{z}_m^t, z_{m+1}] \\ \boldsymbol{y}_{m+1}^t = [\boldsymbol{y}_m^t, y_{m+1}] \end{array}\right.$$

Matrices d'innovation périphériques (MIP)

(En Anglais : Peripheral Innovation matrices PIM)

Une matrice MIP est une matrice qui générée par les trois vecteurs génératrices $\boldsymbol{x}_{m-1}$, $\boldsymbol{z}_{m-1}$, et $\boldsymbol{y}_m$ de façon suivante:

$$\boldsymbol{R} = \left(\begin{array}{ccccccccc} y_1 & \vdots & x_1 & \vdots & x_2 & \vdots & x_3 & \vdots & x_4 \\ \cdots & & & & & & & & \\ z_1 & & y_2 & \vdots & x_1 & \vdots & x_2 & \vdots & x_3 \\ \cdots & \cdots & \cdots & & & & & & \\ z_2 & & z_1 & & y_3 & \vdots & x_1 & \vdots & x_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots & & & & \\ z_3 & & z_2 & & z_1 & & y_4 & \vdots & x_1 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & & \\ z_4 & & z_3 & & z_2 & & z_1 & & y_5 \end{array}\right), \left\{\begin{array}{l} \boldsymbol{x}_{m+1}^t = [x_{m+1}, \boldsymbol{x}_m^t] \\ \boldsymbol{z}_{m+1}^t = [z_{m+1}, \boldsymbol{z}_m^t] \\ \boldsymbol{y}_{m+1}^t = [y_{m+1}, \boldsymbol{y}_m^t] \end{array}\right.$$

Les matrices du types $\boldsymbol{T}$ (Toeplitz), $\boldsymbol{C}$ (circulante), et $\boldsymbol{H}$ (Hankel) sont des cas particuliers des matrices MIP. L'inversion des matrices du type MID et MIP demande un coût de calcul de l'ordre de $O(N_2)$.

Matrices de Vandermonde

Matrice de Vandermonde est une matrice dans laquelle les éléments de chaque colonne sont constitués par la puissance n-ième d'un scalaire. Cette matrice est entièrement définie par le vecteur $\boldsymbol{x} = [x_1, \cdots, x_n]$.

$$\boldsymbol{V} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ x_1^2 & x_2^2 & \cdots & x_n^2 \\ \cdots & \cdots & \cdots & \cdots \\ x_1^{m-1} & x_2^{m-1} & \cdots & x_n^{m-1} \end{pmatrix}$$

$$\det[\boldsymbol{V}] = \prod_{i \neq j} (x_j - x_i)$$

E.3 Matrices blocs

Une matrice quelconque peut toujours être décomposée en blocs. Si ces blocs ont tous les mêmes dimensions on appellera la matrice *matrice-bloc*.
Soit $\boldsymbol{R} = [\boldsymbol{R}_{IJ}]$ une matrice-bloc dont chaque bloc est une matrice de dimensions $[P, P]$. Alors $\boldsymbol{R}$ est:

bloc-symétrique	si	$\boldsymbol{R}_{IJ} = \boldsymbol{R}_{JI}$
bloc-antisymétrique	si	$\boldsymbol{R}_{IJ} = -\boldsymbol{R}_{JI}$
bloc-Hankel	si	$\boldsymbol{R}_{IJ} = \boldsymbol{R}_{I+J}$
bloc-Toeplitz	si	$\boldsymbol{R}_{IJ} = \boldsymbol{R}_{I-J}$

Tous ce que l'on peut dire sur les matrices sont valables sur les matrices-bloc si tous les blocs sont inversibles. Par exemple:
Soit $\boldsymbol{R} = [\boldsymbol{R}_{IJ}]$, $\boldsymbol{I}$, $j = 1, \cdots, N$ une matrice-bloc avec $\boldsymbol{R}_{IJ}$ des matrices de dimensions $[P, P]$. Si toutes les $\boldsymbol{R}_{IJ}$ sont inversibles, alors $\boldsymbol{R}$ peut être factorisée soit par $\boldsymbol{R} = \boldsymbol{L}\boldsymbol{D}\boldsymbol{L}^t$, soit par $\boldsymbol{U}\boldsymbol{R}\boldsymbol{U}^t = \boldsymbol{D}$ où $\boldsymbol{L}$ est une matrice bloc-triangulaire inférieure avec des matrices d'identités sur son diagonale, $\boldsymbol{U}$ est une matrice bloc-triangulaire supérieure, $\boldsymbol{D}$ est une matrice bloc-diagonale, et $\boldsymbol{U} = \boldsymbol{L}^{-1}$.

Inversion d'une matrice décomposée en blocs

Soit $\boldsymbol{A}$ une matrice quelconque telle que

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{pmatrix}$$

Si $\boldsymbol{A}_{22}^{-1}$ existe, alors $\boldsymbol{A}$ peut être factorisée de la façon suivante:

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{I} & \boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1} \\ 0 & \boldsymbol{I} \end{pmatrix} \begin{pmatrix} \boldsymbol{A}_{11} - \boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21} & 0 \\ 0 & \boldsymbol{A}_{22} \end{pmatrix} \begin{pmatrix} \boldsymbol{I} & 0 \\ \boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21} & \boldsymbol{I} \end{pmatrix}$$

On a aussi

$$\operatorname{rang}[\boldsymbol{A}] = \operatorname{rang}\left[\boldsymbol{A}_{11} - \boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}\right] + \operatorname{rang}[\boldsymbol{A}_{22}]$$

$\boldsymbol{A}^{-1}$ existe ssi l'inverse de la matrice $\boldsymbol{T} = \boldsymbol{A}11 - \boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}$ existe. Dans ce cas si on note par $\boldsymbol{B} = \boldsymbol{A}^{-1}$ on a:

$$\boldsymbol{B} = \begin{pmatrix} \boldsymbol{B}_{11} & \boldsymbol{B}_{12} \\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22} \end{pmatrix}$$

$$\boldsymbol{B}_{11} = \boldsymbol{L}^{-1} = (\boldsymbol{A}_{11} - \boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21})^{-1} = \boldsymbol{A}_{11}^{-1} + \boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}\boldsymbol{B}_{22}\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}$$

$$\boldsymbol{B}_{22} = \boldsymbol{D}^{-1} = (\boldsymbol{A}_{22} - \boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12})^{-1} = \boldsymbol{A}_{22}^{-1} + \boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}\boldsymbol{B}_{11}\boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}$$

$$\boldsymbol{B}_{12} = -\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}\boldsymbol{B}_{22} = -\boldsymbol{B}_{11}\boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}$$

$$\boldsymbol{B}_{21} = -\boldsymbol{A}_{22}\boldsymbol{A}_{21}\boldsymbol{B}_{11} = -\boldsymbol{B}_{22}\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}$$

Les matrices $\boldsymbol{L}$ et $\boldsymbol{D}$ sont appelées *le complément Shur* de la matrice $\boldsymbol{A}$.

Cas particulier 1 :
Si $\boldsymbol{A}$ est une matrice bloc triangulaire inférieure, c'est à dire si $\boldsymbol{A}_{21} = 0$, $\boldsymbol{B}$ est aussi triangulaire inférieure, c'est à dire $\boldsymbol{B}_{21} = 0$ et on a:

$$\boldsymbol{B}_{11} = \boldsymbol{A}_{11}^{-1}, \quad \boldsymbol{B}_{22} = \boldsymbol{A}_{22}^{-1}, \quad \boldsymbol{B}_{12} = -\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}$$

Cas particulier 2 :
Si $\boldsymbol{A}$ est une matrice bloc triangulaire supérieure, c'est à dire si $\boldsymbol{A}_{12} = 0$, $\boldsymbol{B}$ est aussi triangulaire supérieure, c'est à dire $\boldsymbol{B}_{12} = 0$ et on a:

$$\boldsymbol{B}_{11} = \boldsymbol{A}_{11}^{-1}, \quad \boldsymbol{B}_{22} = \boldsymbol{A}_{22}^{-1}, \quad \boldsymbol{B}_{21} = -\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}$$

Cas particulier 3 :
Si $\boldsymbol{A}$ est une matrice bloc-diagonale, c'est à dire si $\boldsymbol{A}_{12} = \boldsymbol{A}_{21} = 0$, $\boldsymbol{B}$ est aussi bloc-diagonale, c'est à dire $\boldsymbol{B}_{12} = \boldsymbol{B}_{21} = 0$ et on a:

$$\boldsymbol{B}_{11} = \boldsymbol{A}_{11}^{-1}, \quad \boldsymbol{B}_{22} = \boldsymbol{A}_{22}^{-1}$$

Cas particulier 4 :
Si $\boldsymbol{A}_{22}$ est un scalaire et $\boldsymbol{A}_{21}$ et $\boldsymbol{A}_{12}$ sont des vecteurs, on a:

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{x} \\ \boldsymbol{z}^t & y \end{pmatrix}, \quad \boldsymbol{B} = \boldsymbol{A}^{-1} = \frac{1}{\alpha}\begin{pmatrix} \alpha\boldsymbol{A}_{11}^{-1} + \boldsymbol{w}\boldsymbol{v}^t & \boldsymbol{w} \\ \boldsymbol{v}^t & 1 \end{pmatrix}$$

où α, $\boldsymbol{w}$, et $\boldsymbol{v}$ s'obtient par:

$$\alpha = \frac{1}{(y - \boldsymbol{z}^t\boldsymbol{A}_{11}^{-1}\boldsymbol{x})} = \frac{|\boldsymbol{A}|}{|\boldsymbol{A}_{11}|}, \quad \boldsymbol{w} = -\boldsymbol{A}_{11}^{-1}\boldsymbol{x}, \quad \boldsymbol{v} = -\boldsymbol{A}_{11}^{-t}\boldsymbol{z}$$

E.4 Matrices de Toeplitz et de Hankel

E.4.1 Matrices de Toeplitz

Une matrice de dimensions $[M, N]$ est dite de Toeplitz si

$$[\boldsymbol{T}]_{ij} = t_{(i-j)}, \quad i = 1, \cdots, M, \quad j = 1, \cdots, N$$

$$\boldsymbol{T} = \begin{pmatrix} t_0 & t_{-1} & . & . & t_{-(N-2)} & t_{(N-1)} \\ t_1 & t_0 & . & . & . & t_{-(N-2)} \\ . & . & . & . & . & . \\ . & . & t_0 & . & . & . \\ t_{(M-2)} & . & . & . & t_0 & t_{-1} \\ t_{(M-1)} & t_{(M-2)} & . & . & t_1 & t_0 \end{pmatrix}$$

Dans le cas où $\boldsymbol{T}$ est carrée on a $[\boldsymbol{T}]_{ij} = t_{(i-j)}, \quad i, j = 1, \cdots, N$. Si de plus $\boldsymbol{T}$ est symétrique on a $[\boldsymbol{T}]_{ij} = t_{|i-j|}, \quad i, j = 1, \cdots, N$.

Par la suite nous considérons seulement les matrices de Toeplitz carrées de dimensions $[N, N]$.

$$\boldsymbol{T} = \begin{pmatrix} t_0 & t_{-1} & . & . & t_{-(N-2)} & t_{-(N-1)} \\ t_1 & t_0 & . & . & . & t_{-(N-2)} \\ . & . & . & . & . & . \\ . & . & t_0 & . & . & . \\ t_{(N-2)} & . & . & . & t_0 & t_{-1} \\ t_{(N-1)} & t_{(N-2)} & . & . & t_1 & t_0 \end{pmatrix}$$

Une matrice de Toeplitz est entièrement déterminée par sa première ligne

$$[\boldsymbol{T}]_{1*} = [t_0, t_{-1}, \cdots, t_{-(N-2)}, t_{-(N-1)}]^t$$

et sa première colonne

$$[\boldsymbol{T}]_{*1} = [t_0, t_1, \cdots, t_{(N-2)}, t_N - 1]^t$$

c'est à dire par $(2N - 1)$ éléments $\{t_{-(N-1)}, \cdots, t_0, \cdots, t_{(N-1)}\}$.

Si $\boldsymbol{T}_N$ est une matrice de Toeplitz de dimensions $[N, N]$ et si on note par $\{\boldsymbol{T}_n, \quad n = 1, \cdots, N\}$ les sous-matrices de $\boldsymbol{T}_N$ on a les relations suivantes:

$$\boldsymbol{J}_n \boldsymbol{T}_n \boldsymbol{J}_n = \boldsymbol{T}_n^t, \quad n = 1, \cdots, N$$

où $\boldsymbol{J}_n$ est une matrice d'inversion d'ordre de dimensions $[n, n]$. Une matrice de Toeplitz peut être factorisée de façon suivante:

$$\boldsymbol{T} = \boldsymbol{L}\boldsymbol{U}$$

De plus si elle est hermitienne elle peut être factorisée de façon suivante:

$$\boldsymbol{T} = \boldsymbol{L}\boldsymbol{D}\boldsymbol{L}^*$$

Les lignes successives de $\boldsymbol{L}$ sont les coefficients des prédicateurs linéaires d'ordre $0, 1, \cdots, N-1$, et les éléments diagonaux de la matrice $\boldsymbol{D}$ sont les variances d'erreurs de prédictions correspondantes.

Si $\boldsymbol{T}$ est une matrice de Toeplitz, on a:

$$\boldsymbol{T}^{-1} = \frac{1}{\lambda_{n-1}} \begin{pmatrix} 1 & \boldsymbol{b}_{n-1}^t \\ \boldsymbol{a}_{n-1} & \boldsymbol{S}_{n-1} \end{pmatrix}$$

Les éléments de $\boldsymbol{S}_{n-1}$ sont déterminés complètement par λ_{n-1}, $\boldsymbol{a}_{n-1}$, et $\boldsymbol{b}_{n-1}$.

Une décomposition moins connue est:

$$\boldsymbol{T}^{-1} = \frac{1}{\lambda_{n-1}}[\boldsymbol{A}!\boldsymbol{B}! - \boldsymbol{B}\boldsymbol{A}]$$

où $\boldsymbol{A}!$ et $\boldsymbol{B}$ sont des matrices Toeplitz triangulaire supérieur type $\boldsymbol{L}_1$, et $\boldsymbol{A}$ et $\boldsymbol{B}!$ sont des matrices matrices Toeplitz triangulaire inférieur type $\boldsymbol{U}_1$. Les éléments de ces matrices sont données par:

$$[\boldsymbol{A}]_{jk} = [\widetilde{\boldsymbol{a}}_{n-1}]_{k-j}, \quad [\boldsymbol{B}]_{jk} = [\widetilde{\boldsymbol{b}}_{n-1}]_{k-j}$$

$$[\boldsymbol{A}!]_{jk} = [\boldsymbol{a}_{n-1}]_{j-k} + \delta_{j-k}, \quad [\boldsymbol{B}!]_{jk} = [\boldsymbol{b}_n - 1]_{kj} + \delta_{j-k}, \quad 1 \leq j, k \leq n$$

et où pour tout vecteur $\boldsymbol{g}_m$, $\quad \boldsymbol{g}_m(k) = 0$, $\forall k \leq 0$ ou $k > m$.

E.4.2 Matrices de Hankel

$\boldsymbol{H}$ est une matrice de Hankel si tous ses éléments se trouvant sur des lignes parallèle à sa deuxième diagonale sont identiques. Par exemple :

$$\boldsymbol{H} = \begin{pmatrix} h_N & h_{-N+1} & . & . & h_{-1} & h_0 \\ h_{-N+1} & . & . & . & h_0 & h_1 \\ \vdots & . & . & . & . & \vdots \\ h_{-1} & h_0 & . & . & . & h_{M-1} \\ h_0 & h_1 & . & . & & h_M \end{pmatrix}$$

Si $\boldsymbol{H}$ est une matrice carrée on la note par $\boldsymbol{H}_N$.

$\boldsymbol{H}_N$ est une matrice de Hankel d'ordre N si

$$[\boldsymbol{H}]_{ij} = h_{i+j-2}, \quad i, j = 1, \cdots, N$$

$$\boldsymbol{T} = \begin{pmatrix} h_0 & h_1 & . & . & h_{N-2} & h_{N-1} \\ h_1 & h_2 & . & . & h_{N-1} & h_N \\ . & . & . & . & . & . \\ . & . & h_{N-1} & . & . & h_{2N-3} \\ . & . & . & . & h_{2N-2} & h_{2N-2} \\ h_N & h_{N+1} & & . & h_{2N-2} & h_{2N-1} \end{pmatrix}$$

Une matrice de Hankel est toujours symétrique: $\boldsymbol{H}^t = \boldsymbol{H}$. Elle est hermitienne ssi tous ses éléments sont réels. Si T est une matrice de Toeplitz, alors les matrices

$$\boldsymbol{H}_1 = \boldsymbol{T}\boldsymbol{J}, \quad \text{et} \quad \boldsymbol{H}_2 = \boldsymbol{J}\boldsymbol{T}^t$$

sont des matrices de Hankel.

Les trois systèmes d'équations suivantes sont équivalentes:

$$\boldsymbol{H}\boldsymbol{a} = -\boldsymbol{d}, \quad \boldsymbol{T}[\boldsymbol{J}\boldsymbol{a}] = -\boldsymbol{d}, \quad \boldsymbol{T}\boldsymbol{a} = -\boldsymbol{J}\boldsymbol{d}$$

E.5 Matrices circulantes

Une matrice circulante est une matrice qui est définie par seule sa première ligne (ou seule par sa première colonne). $\boldsymbol{C}$ est une matrice *circulante à droite (circulante vers le haut)* si :

$$[\boldsymbol{C}]_{ij} = \begin{cases} c_{j-i} & \text{si } \; j-i \geq 0 \\ c_{n-j+1} & \text{si } \; j-i < 0 \end{cases}$$

Exemple :

$$\boldsymbol{C} = \begin{pmatrix} c_0 & c_1 & c_2 & \cdots & \cdots & c_{n-1} \\ c_{n-1} & c_0 & c_1 & \cdots & \cdots & c_{n-2} \\ c_{n-2} & c_{n-1} & c_0 & \cdots & \cdots & . \\ \vdots & c_{n-2} & c_{n-1} & c_0 & \cdots & \vdots \\ . & \cdots & \cdots & \cdots & \cdots & c_1 \\ c_1 & c_2 & \cdots & \cdots & c_{n-1} & c_0 \end{pmatrix}$$

Une matrice circulante à droite est une matrice de Toeplitz qui est symétrique par rapport à sa deuxième diagonale. $\boldsymbol{C}$ est une matrice *circulante à gauche (ou circulante vers le bas)* si :

$$[\boldsymbol{C}]_{ij} = \begin{cases} c_{n+1-i-j} & \text{si } \; j+i \leq n+1 \\ c_{2n+1-i-j} & \text{si } \; j+i > n+1 \end{cases}$$

Exemple :

$$\boldsymbol{C} = \begin{pmatrix} c_{n-1} & \cdots & \cdots & c_2 & c_1 & c_0 \\ c_{n-2} & c_{n-1} & \cdots & \cdots & c_0 & c_{n-1} \\ c_{n-3} & \cdots & \cdots & c_0 & c_{n-1} & c_{n-2} \\ \vdots & \cdots & c_0 & c_{n-1} & c_{n-2} & \vdots \\ c_1 & \cdots & \cdots & \cdots & \cdots & c_2 \\ c_0 & c_{n-1} & c_{n-2} & \cdots & c_2 & c_1 \end{pmatrix}$$

Une matrice circulante à gauche est une matrice de Hankel qui est symétrique par rapport à sa diagonale. Si $\boldsymbol{C}$ est une matrice circulante de dimensions $[N, N]$ avec des éléments réels $[\boldsymbol{C}]_{ij}$ on a:

$$[\boldsymbol{C}]_{ij} = c_{(i-j) \bmod N}, \quad \boldsymbol{C}^t = \boldsymbol{C}^{-1}$$

$$\boldsymbol{C} = \begin{pmatrix} c_0 & c_{N-1} & . & . & c_2 & c_1 \\ c_1 & c_0 & . & . & . & c_2 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ c_{N-2} & . & . & . & c_0 & c_{N-1} \\ c_{N-1} & c_{N-2} & . & . & c_1 & c_0 \end{pmatrix}$$

Les valeurs propres d'une matrice circulante sont les coefficients de la TFD de sa première ligne et ses vecteurs propres sont les vecteurs de base de la TFD.

Si $\boldsymbol{C}$ est une matrice circulante, alors

$$\boldsymbol{C} = \boldsymbol{F}\boldsymbol{\Lambda}\boldsymbol{F}^* = \boldsymbol{F}^*\boldsymbol{\Lambda}_c^*\boldsymbol{F}$$

où

- $\boldsymbol{F}$ est la matrice de la TFD :

$$\boldsymbol{F} = \frac{1}{\sqrt{N}}[\boldsymbol{w}_0, \boldsymbol{w}_1, \cdots, \boldsymbol{w}_p, \cdots, \boldsymbol{w}_{n-1}]$$

$$[\boldsymbol{w}_p]_q = \exp\left[-j\frac{2\pi pq}{N}\right], \quad p = 0, \cdots, N-1 \quad q = 1, \cdots, N$$

- $\boldsymbol{\Lambda}_c$ est une matrice diagonale contenant les valeurs propres de la matrice $\boldsymbol{C}$, c'est à dire les coefficients de la TFD de sa première ligne.

- $\boldsymbol{F}$ est une matrice qui multipliée par un vecteur $\boldsymbol{x}$ donne un vecteur $\bar{\boldsymbol{x}}$ qui contient les valeurs de sa TFD.

$$\boldsymbol{F}\boldsymbol{x} = \bar{\boldsymbol{x}} \quad \text{et} \quad \boldsymbol{x} = \boldsymbol{F}^*\bar{\boldsymbol{x}}$$

E.6 Matrices bloc-Toeplitz et bloc-circulantes

E.6.1 Matrices bloc-Toeplitz

La matrice $\boldsymbol{R}$ est dite bloc-Toeplitz si $\boldsymbol{R}$ est une matrice-bloc dont la IJ-ème bloc $\boldsymbol{R}_{IJ}$ est fonction de $I-J$):

$$\boldsymbol{R}_{IJ} = \boldsymbol{R}_{I-J}.$$

Si chaque bloc est aussi une matrice de Toeplitz alors $\boldsymbol{R}$ est une matrice Toeplitz-bloc-Toeplitz.

Matrices bloc-Toeplitz avec des blocs Toeplitz

Soit $\boldsymbol{R}_{P,M}$ est une matrice bloc d'ordre M avec les blocs $\boldsymbol{R}_{IJ}, \quad I, J = 1, \cdots, M$, qui aux même sont des matrices de Toeplitz de dimensions $[P, P]$, c'est à dire:

$$\boldsymbol{R}_{PM} = \begin{pmatrix} R_0 & R_{-1} & . & . & R_{-(M-2)} & R_{-(M-1)} \\ R_1 & R_0 & . & . & . & R_{-(M-2)} \\ . & . & . & . & . & . \\ . & . & R_0 & . & . & . \\ R_{(M-2)} & . & . & . & R_0 & R_{-1} \\ R_{(M-1)} & R_{(M-2)} & . & . & R_1 & R_0 \end{pmatrix}$$

avec

$$\boldsymbol{R}_J = \begin{pmatrix} r_0 & r_{-1} & . & . & r_{-(P-2)} & r_{-(P-1)} \\ r_1 & r_0 & . & . & . & r_{-(P-2)} \\ . & . & . & . & . & . \\ . & . & r_0 & . & . & . \\ r_{(P-2)} & . & . & . & r_0 & r_{-1} \\ r_{(P-1)} & r_{(P-2)} & . & . & r_1 & r_0 \end{pmatrix}$$

Soit $\boldsymbol{J}_{P,M}$ une matrice de dimensions $[PM, PM]$ telle que

$$\boldsymbol{J}_{PM} = \begin{pmatrix} . & . & . & . & \boldsymbol{J}_P \\ . & 0 & . & \boldsymbol{J}_P & . \\ . & . & \boldsymbol{J}_P & . & . \\ . & \boldsymbol{J}_P & . & 0 & . \\ \boldsymbol{J}_P & . & . & . & . \end{pmatrix} \quad \text{avec} \quad \boldsymbol{J}_P = \begin{pmatrix} . & . & . & . & 1 \\ . & 0 & . & 1 & . \\ . & . & 1 & . & . \\ . & 1 & . & 0 & . \\ 1 & . & . & . & . \end{pmatrix}$$

une matrice de dimensions $[P, P]$. Alors on a les propriétés suivantes:

- $\boldsymbol{R}_{P,M}$ et son inverse sont toutes les deux des matrices centrosymétriques et on a

$$\boldsymbol{J}_{P,M}\boldsymbol{R}_{P,M}\boldsymbol{J}_{P,M} = \boldsymbol{R}_{P,M}^t \quad \text{et} \quad \boldsymbol{J}_{P,M}\boldsymbol{R}_{P,M}^{-1}\boldsymbol{J}_{P,M} = \boldsymbol{R}_{P,M}^{-t}$$

- $\boldsymbol{R}_{P,M}$ qui est une matrice $\boldsymbol{TT}$ de $[M, M]$ blocs de matrice de Toeplitz de dimensions $[P, P]$ peut être transformée à une matrice $\boldsymbol{TT}$ de $[P, P]$ blocs de matrices de Toeplitz de dimensions $[M, M]$ avec de simples permutations de ses éléments. Ceci peut être démontré en définissant une matrice de permutation $\boldsymbol{E}$ de dimensions $[PM, PM]$ telle que

$$\boldsymbol{E} = (\boldsymbol{E}_0 \quad . \quad . \quad \boldsymbol{E}_{P-1})$$

avec $\boldsymbol{E}_j$ une matrice de dimensions $[M, MP]$ telle que

$$\boldsymbol{E}_j = \begin{pmatrix} \boldsymbol{e}_i & . & . & . & . \\ . & \boldsymbol{e}_i & . & 0 & . \\ . & . & . & . & . \\ . & 0 & . & \boldsymbol{e}_i & . \\ . & . & . & . & \boldsymbol{e}_i \end{pmatrix}$$

avec $\boldsymbol{e}_i$ un vecteur ligne dont toutes ses composantes sont nulles sauf la j-ème qui est égale à 1. On montre alors que

$$\boldsymbol{E}\boldsymbol{R}_{P,M}\boldsymbol{E}^t = \widetilde{\boldsymbol{R}}_{M,P}$$

où $\widetilde{\boldsymbol{R}}_{M,P}$ est une matrice $\boldsymbol{TT}$ de $[PxP]$ blocs de matrices de Toeplitz de dimensions $[M, M]$. La ij-ème bloc de $\widetilde{\boldsymbol{R}}_{M,P}$ s'obtient par $\boldsymbol{E}_j\boldsymbol{R}_{P,M}\ \boldsymbol{E}_j^t$ et le kl-ème élément de ce bloc est $\boldsymbol{e}_j\boldsymbol{R}_{l-k}\boldsymbol{e}_j^t = r_{j-i,l-k}$.

$$\boldsymbol{E}\boldsymbol{R}_{P,M}\boldsymbol{E}^t = \begin{pmatrix} \boldsymbol{E}_0 \\ . \\ . \\ \boldsymbol{E}_{P-1} \end{pmatrix} \boldsymbol{R}_{P,M}[\boldsymbol{E}^t, \cdots, \boldsymbol{E}_{p-1}^t]$$

E.6.2 Matrices bloc-circulantes

Si $\boldsymbol{A}$ est une matrice bloc-circulante, composée de $[P, P]$ blocs de matrices circulante de dimensions $[K, K]$, $\boldsymbol{A}$ peut être factorisée de façon suivante:

$$\boldsymbol{A} = \boldsymbol{F}^*\boldsymbol{D}\boldsymbol{F}$$

$$\boldsymbol{F} = \begin{pmatrix} F & . & . & . & . \\ . & F & . & 0 & . \\ . & . & . & . & . \\ 0 & . & F & . & . \\ . & . & . & . & F \end{pmatrix} D = \begin{pmatrix} D_{11} & . & . & . & D_{1p} \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ D_{P1} & . & . & . & D_{PP} \end{pmatrix}$$

Chaque matrice D_{ij} est une matrice diagonale contenant les valeurs propres de la ij-ème bloc (matrice circulante) de $\boldsymbol{A}$.

E.7 Inversion récursive des matrices

Dans cette section nous allons présenter une méthode d'inversion récursive des matrices qui est basée sur le partitionnement de matrices en blocs.

Cas général

Si $\boldsymbol{A}_n$ est une matrice définie positive d'ordre n, alors on peut la partionner de façon suivante:

$$\boldsymbol{A}_n = \begin{pmatrix} \boldsymbol{A}_{n-1} & \boldsymbol{a}_n \\ \boldsymbol{b}_n^t & a_{nn} \end{pmatrix}$$

où $\boldsymbol{a}_n$ et $\boldsymbol{b}_n$ sont définie par

$$\boldsymbol{a}_n = [a_{1n}, a_{2n}, \cdots, a_{n-1,n}]^t \quad \text{et} \quad \boldsymbol{b}_n = [b_{1n}, n_{2n}, \cdots, b_{n-1,n}]^t$$

Si $\boldsymbol{Q}_n = \boldsymbol{A}_n^{-1}$ est aussi partitionnée par

$$\boldsymbol{Q}_n = \boldsymbol{A}_n^{-1} = \begin{pmatrix} \boldsymbol{Q}_{n-1} & \boldsymbol{Q}_n \\ \boldsymbol{v}_n^t & q_{nn} \end{pmatrix}$$

on a

$$\boldsymbol{Q}_{n-1} = \boldsymbol{A}_{n-1}^{-1} + \alpha_n \boldsymbol{Q}_n \boldsymbol{v}_n^t$$

où

$$\alpha_n = a_{nn} - \boldsymbol{b}_n \boldsymbol{A}_{n-1}^{-1} \boldsymbol{a}_n$$

$$q_{nn} = \frac{1}{\alpha_n}, \quad \boldsymbol{Q}_n = -\frac{1}{\alpha_n}[\boldsymbol{A}_{n-1}^{-1}\boldsymbol{a}_n], \quad \boldsymbol{v}_n = \frac{1}{\alpha_n}[\boldsymbol{A}_{n-1}^{-1}\boldsymbol{b}_n].$$

Cas des matrices symétriques

Si $\boldsymbol{A}$ est une matrice réelle, symétrique et définie positive, alors son inverse $\boldsymbol{A}^{-1}$ ainsi que toutes ses sous-matrices principales sont réelles, symétriques et définie positives. Si on note alors par $\boldsymbol{A}_n^{-1} = \boldsymbol{Q}_n$ et par $[\boldsymbol{Q}_n]_{ij} = q_{ij}^{(n)}$ on a

$$q_{nn}^{(n)} = \frac{1}{\left(a_{nn} - \sum_{i=1}^{n-1} a_{in}\beta_i^{(n)}\right)}$$

$$q_{ij}^{(n)} = q_{ij}^{(n-1)} - q_{in}^{(n)}\beta_i^{(n)} = q_{ij}^{(n-1)} - \beta_i^{(n)} q_{jn}^{(n)}$$

$$q_{in}^{(n)} = -q_{in}^{(n)}\beta_i^{(n)}, \quad i = 1, \cdots, n-1, \quad \beta_i^{(n)} = \sum_{k=1}^{n-1} q_{ik}^{(n-1)} a_{kn}$$

avec initialisation $\beta_1^{(1)} = -1$ et $q_1^{(1)} = \frac{1}{a_{11}}$.

On a aussi la propriétés suivante

$$|\boldsymbol{A}_n| = \prod_{i=1}^{N} \frac{1}{q_{ii}^{(i)}} = \frac{1}{q_{11}^{(1)} q_{22}^{(2)} \cdots q_{nn}^{(n)}}$$

Cas des matrices de Toeplitz symétriques

$$\beta_n = \sum_{k=1}^{n-1} q_{1k}^{(n-1)} a_{1,n+1-k}, \quad q_{11}^{(n)} = \frac{q_{11}^{(n-1)}}{1-\beta_n^2}, \quad q_{1n}^{(n)} = -\beta_n q_{11}^{(n)}$$

Si $\boldsymbol{A}$ est une matrice définie positive, alors on montre que $0 \leq \beta_n < 1$, et que $q_{11}^{(n)}$ croit d'une façon monotone en fonction de n, c'est à dire que $(q_{11}^{(n)}/q_{11}^{(n-1)}) \geq 1$. D'autre part si β_n devient nul à l'ordre n, alors $q_{11}^{(n)}$ devient aussi nul, et les éléments de $n+1$ème inverse ne changent plus.

Cas des matrices de corrélation d'un signal stationnaire

La matrice de corrélation d'un signal stationnaire exponentiellement est une matrice de Toeplitz symétrique avec la propriété suivante

$$\frac{a_{12}}{a_{11}} = \frac{a_{13}}{a_{12}} = \cdots = \frac{a_{1N}}{a_{N-1}}$$

Dans ce cas on montre que l'on peut écrire

$$\beta_2 = \frac{a_{12}}{a_{11}}, \quad \beta_k = 0 \quad \text{pour} \quad k > 2,$$

$$q_{11}^{(n)} = q_{nn}^{(n)} = \frac{1}{a_{11}(1-\beta_2^2)}, \quad q_{12}^{(n)} = q_{21}^{(n)} = \frac{-\beta_2}{a_{11}(1-\beta_2^2)}$$

$$q_{jj}^{(n)} = \frac{1+\beta_2^2}{a_{11}(1-\beta_2^2)}, \quad j = 2, \cdots, n-1, \quad q_{ij}(n) = 0 \quad |i-j| > 2$$

D'une façon plus schématique $\boldsymbol{A}_n^{-1}$ est sous la forme de

$$\boldsymbol{A}_n^{-1} = \frac{1}{a_{11}(1-\beta_2^2)} \begin{pmatrix} 1 & -\beta_2 & . & . & . \\ -\beta_2 & (1+\beta_2^2) & . & 0 & . \\ . & . & . & . & . \\ . & 0 & . & (1+\beta_2^2) & -\beta_2 \\ . & . & . & -\beta_2 & 1 \end{pmatrix}$$

Cas des matrices de Toeplitz symétriques et tridiagonale

Un cas particulier des matrices de Toeplitz symétrique est une matrice dont les éléments du m-ième diagonal sont tous égale à zéro pour m supérieure à k. Pour $k = 2$ on a une matrice tridiagonale. Ceci est par exemple le cas de la matrice de covariance d'un processus AR d'ordre 1.

$$\boldsymbol{A}_n = \begin{pmatrix} a_{11} & a_{12} & 0 & . & 0 \\ a_{12} & a_{11} & . & 0 & . \\ 0 & . & . & . & 0 \\ . & 0 & . & a_{11} & \\ & 0 & . & a_{12} & a_{11} \end{pmatrix}$$

Pour cette matrice la valeur de β_n devient

$$\beta_n = \frac{(-1)^n (a_{12}/a_{11})^{n-1}}{\prod_{j=2}^{n-1}(1-\beta_j^2)^{n-j}} = (-1)^n a_{12}^{n-1} \det\left[\boldsymbol{A}_{n-1}^{-1}\right], \quad n = 2, 3, \cdots$$

Une propriétés intéressante de ces matrices est le faite que même si la matrice est définie positive à l'ordre n, elle peut ne pas l'être à l'ordre n+1. Ceci peut être le cas si par exemple $|\beta_{n+1}| \geq 1$. On peut montrer que si $|\beta_2| \leq 1/2$, alors $|\beta_n| \longrightarrow 0$ quand $n \longrightarrow \infty$. Ainsi la matrice reste définie positive pour toute valeur de n, et si $|\beta_2| > 1/2$, la matrice deviendra non définie positive.

Cas des matrices triangulaires

Si $\boldsymbol{A}$ est une matrice réelle, triangulaire et définie positive d'ordre n, alors son inverse $\boldsymbol{A}_n^{-1}$ ainsi que toutes ses sous-matrices principales sont réelles, triangulaires et définie positives. Ceci est un cas particulier de (1.1) et (1.2) avec $b_n = 0$ et $\boldsymbol{v}_n = 0$. Ainsi les éléments q_{ij} de la matrice $\boldsymbol{Q}_n$ sont données par

$$q_{ij} = \begin{cases} -\dfrac{1}{a_{ii}} \displaystyle\sum_{k=1}^{j-1} q_{ik} a_{kj} & \text{pour} \quad j > i \\ \dfrac{1}{a_{ii}} & \text{pour} \quad j = i \\ 0 & \text{pour} \quad j < i \end{cases}$$

Cas de Toeplitz et triangulaires

Si $\boldsymbol{A}$ est une matrice de Toeplitz, réelle, triangulaire et définie positive d'ordre n, alors son inverse $\boldsymbol{A}_n^{-1}$ ainsi que toutes ses sous- matrices principales sont des matrices de Toeplitz, réelles, triangulaires et définie positives.

$$\boldsymbol{A}_n = \begin{pmatrix} a_0 & a_1 & . & . & a_{n-2} & a_{n-1} \\ 0 & a_0 & a_1 & . & & a_{n-2} \\ . & 0 & . & . & & \\ . & . & . & . & & . \\ . & . & . & . & & . \\ . & 0 & & & a_0 & a_1 \\ . & . & . & 0 & & a_0 \end{pmatrix}$$

Dans ce cas on a $a_{ij} = a_{i-j}, \quad q_{ij} = q_{i-j} = 0$ pour $i > j$, et on a $q_0 = 1/a_0$, et

$$q_{ij} = -\frac{1}{a_0} \sum_{k=1}^{j-1} q_{k-i} a_{j-k}$$

Si on note par $m = j - i$ et $n = k - i$, alors on a

$$q_m = -\frac{1}{a_0} \sum_{n=1}^{m-1} q_n a_{m-n}$$

$$\boldsymbol{Q}_n = \boldsymbol{A}_n^{-1} = \begin{pmatrix} q_0 & q_1 & . & . & q_{n-2} & q_{n-1} \\ 0 & q_0 & q_1 & . & . & q_{n-2} \\ . & 0 & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & 0 & q_0 & q_1 \\ 0 & . & . & . & 0 & q_0 \end{pmatrix}$$

E.8 Résolution des systèmes d'équations linéaires

Dans cette section nous allons présenter les deux problèmes essentiels du calcul matriciel qui sont la résolution d'un système linéaire et le calcul des valeurs propres et des vecteurs propres d'une matrice. L'objective étant de mettre en évidence les difficultés de ces deux problèmes lors que la matrice est mal conditionnée.

Les deux problèmes fondamentaux de l'analyse numérique matricielles sont:

1. Résolution d'un système linéaire $\boldsymbol{A}\boldsymbol{u} = \boldsymbol{b}$.

2. Calcul des valeurs propres et des vecteurs propres d'une matrice $\boldsymbol{A}$.

Pour résoudre un problème du type (il) ou (ii), on commence naturellement par choisir une méthode déterminée (selon des critères qui seront précisés au cours des chapitres suivants). Or, quelle qui soit la méthode, ce résultat n'est pas en général le résultat exacte, car certaines erreurs sont inévitables au cours du calcul ce sont les erreurs d'arrondi.

Alors que les erreurs d'arrondi sont toujours présentes dans un calcul, le deuxième type d'erreurs n'apparaît que dans certaines méthodes. Une méthode directe conduit à la solution exacte du problème en un nombre fini d'opérations élémentaires s'il n'y avait pas d'erreurs d'arrondi.

Dans une méthode récursive la solution $\boldsymbol{u}$ du problème est la limite d'une suite $\boldsymbol{u}_k$ de solutions exactes des sous-matrices de $\boldsymbol{A}$. Dans une méthode itérative la solution $\boldsymbol{u}$ du problème est la limite d'une suite $\boldsymbol{u}_k$ de solutions approchées. Comme on est bien obligé d'arrêter les calcul à l'itération k on commet ainsi une erreur de troncature mesurée par le nombre $\|\boldsymbol{u} - \boldsymbol{u}_k\|$ (pour une norme vectorielle donnée). Le choix d'une méthode dépends des propriétés de la matrice : matrice pleine quelconque matrice creuse; matrice symétrique; matrice-bandes; etc.

Une des premières caractéristiques de la matrice à considérer est le nombre et, éventuellement la répartition des éléments nuls de la matrice. C'est ainsi que l'on distingue les matrices pleines et le matrices creuses. Les problème de résolution de système linéaire les plus faciles à traiter numériquement sont ceux dont les matrices sont symétriques définies positives.

Les problème de calcul des valeurs propres et des vecteurs propres d'une matrice les plus faciles à traiter numériquement sont ceux des matrices sont symétriques. Autant il existe les trois types de méthodes directe, récursive et itérative pour le problème (1), les méthodes de calcul des valeurs propres ne sauraient qu'être itératives. En effet, les valeurs propres d'une matrice sont les racines d'un polynôme de degré n. Or l'existence d'une méthode directe pour le calcul des valeurs propres équivaudrait à affirmer qu'on peut calculer les racines d'un polynôme arbitraire en un nombre fini d'opérations élémentaires.

E.8.1 Conditionnement d'un problème de résolution des systèmes linéaires

Considérons la matrice inversible

$$\boldsymbol{A} = \begin{pmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{pmatrix}$$

et son inverse

$$\boldsymbol{A}^{-1} = \begin{pmatrix} 25 & -41 & 10 & -6 \\ -41 & 68 & -17 & 10 \\ 10 & -17 & 5 & -3 \\ -6 & 10 & -3 & 2 \end{pmatrix}$$

et les deux vecteurs $\boldsymbol{x}^t = [x_1, \cdots, x_4]$ et $\boldsymbol{y}^t = [32, 23, 33, 31]$. La solution exacte de l'équation $\boldsymbol{Ax} = \boldsymbol{y}$ est $\boldsymbol{x}^t = [1, 1, 1, 1]$. Considérons maintenant le vecteur perturbé $[\boldsymbol{y} + \delta \boldsymbol{y}]^t = [32.1, 22.9, 33.1, 30.9]$ très légèrement différent à $\boldsymbol{y}$. La solution de l'équation $\boldsymbol{Ax} = [\boldsymbol{y} + \delta\boldsymbol{y}]$ est

$$\boldsymbol{x}^t = [9.2, -12.6, 4.5, -1.1]$$

Autrement dit une erreur relative de 1/200 sur les données $\boldsymbol{y}$ entraîne une erreur relative de 10/1 sur le résultat (soit un rapport d'amplification des erreurs relatives de l'ordre de 2000) de la solution du système linéaire. Considérons également une matrice $\boldsymbol{A}$ légèrement modifiée :

$$\boldsymbol{A} = \begin{pmatrix} 10 & 7 & 8.1 & 7.2 \\ 7.08 & 5.04 & 6 & 5 \\ 8 & 5.98 & 9.89 & 9 \\ 6.99 & 4.99 & 9 & 9.98 \end{pmatrix}$$

cette fois ci encore la solution de l'équation $\boldsymbol{Ax} = \boldsymbol{y}$ est $\boldsymbol{x}^t = [-81, 137, -34, 22]$. Là encore, de très petites variations des éléments de la matrice modifient complètement le résultat. Pourtant, la matrice $\boldsymbol{A}$ du système a un bon aspect: elle est symétrique, son déterminant vaut 1, et la matrice inverse $\boldsymbol{A}^{-1}$ est tout aussi sympathique ! Ces deux exemples montrent qu'il est utile d'étudier le conditionnement d'une matrice quand on a à résoudre un système d'équations linéaires même si la solution est calculée par une méthode directe. Analyse des erreurs permet d'établir facilement les relations suivantes: soient $\boldsymbol{Ax} = \boldsymbol{y}$, $\boldsymbol{A}(\boldsymbol{x} + \delta\boldsymbol{x}) = \boldsymbol{y} + \delta\boldsymbol{y}$. Alors on a:

$$\frac{||\delta\boldsymbol{x}||}{||\boldsymbol{x}||} \leq ||\boldsymbol{A}|| \; ||\boldsymbol{A}^{-1}|| \frac{||\delta\boldsymbol{y}||}{||\boldsymbol{y}||}$$

soient

$$\boldsymbol{Ax} = \boldsymbol{y}, \quad (\boldsymbol{A} + \Delta\boldsymbol{A})(\boldsymbol{x} + \delta\boldsymbol{x}) = \boldsymbol{y}$$

Alors

$$\frac{||\delta\boldsymbol{x}||}{||\boldsymbol{x} + \delta\boldsymbol{x}||} \leq ||\boldsymbol{A}|| \; ||\boldsymbol{A}^{-1}|| \frac{||\Delta\boldsymbol{A}||}{||\boldsymbol{A}||}$$

On définit ainsi le nombre de conditionnement de la matrice $\boldsymbol{A}$ par

$$\text{cond}\,\{\boldsymbol{A}\} = ||\boldsymbol{A}|| \; ||\boldsymbol{A}^{-1}||$$

Ces deux inégalités montre que le nombre cond $\{\boldsymbol{A}\}$ mesure la sensibilité de la solution $\boldsymbol{x}$ du système linéaire $\boldsymbol{Ax} = \boldsymbol{y}$ vis- à-vis des variations sur les données $\boldsymbol{A}$ et $\boldsymbol{y}$, qualité que l'on appelle le conditionnement du système linéaire considéré. Un système linéaire est bien-ou-mal-conditionné selon que cond $\{\boldsymbol{A}\}$ est petit ou grand. Par la suite nous présentons un certains nombre de théorèmes et résultats (sans démonstrations) concernant cond $\{\boldsymbol{A}\}$.

Théorème 3 *Soit $\boldsymbol{A}$ une matrice inversible, et soit $\boldsymbol{x}$ et $(\boldsymbol{x}+\delta\boldsymbol{x})$ les solutions des systèmes linéaires $\boldsymbol{Ax} = \boldsymbol{y}$ et $\boldsymbol{A}(\boldsymbol{x} + \delta\boldsymbol{x}) = \boldsymbol{y} + \delta\boldsymbol{y}$. Alors l'inégalité*

$$\frac{\|\delta\boldsymbol{x}\|}{\|\boldsymbol{x}\|} \leq cond\{\boldsymbol{A}\} \frac{\|\delta\boldsymbol{y}\|}{\|\boldsymbol{y}\|}$$

est satisfaite, et c'est la meilleure possible— pour une matrice $\boldsymbol{A}$ donnée, on peut trouver des vecteurs $\boldsymbol{y} \neq 0$ et $\delta\boldsymbol{y} \neq 0$ tels qu'elle devienne une égalité.

Théorème 4 *Soit $\boldsymbol{A}$ une matrice inversible, et soit $\boldsymbol{x}$ et $(\boldsymbol{x}+\delta\boldsymbol{x})$ les solutions des systèmes linéaires $\boldsymbol{Ax} = \boldsymbol{y}$ et $(\boldsymbol{A} + \Delta\boldsymbol{A})(\boldsymbol{x} + \delta\boldsymbol{x}) = \boldsymbol{y}$. Alors l'inégalité*

$$\frac{\|\delta\boldsymbol{x}\|}{\|\boldsymbol{x} + \delta\boldsymbol{x}\|} \leq cond\{\boldsymbol{A}\} \frac{\|\Delta\boldsymbol{A}\|}{\|\boldsymbol{A}\|}$$

est satisfaite, et c'est la meilleure possible— pour une matrice $\boldsymbol{A}$ donnée, on peut trouver des vecteurs $\boldsymbol{y} \neq 0$ et une matrice $\Delta\boldsymbol{A} \neq 0$ tels qu'elle devienne une égalité. On a par ailleurs l'inégalité

$$\frac{\|\delta\boldsymbol{x}\|}{\|\boldsymbol{x}\|} \leq cond\{\boldsymbol{A}\} \frac{\|\Delta\boldsymbol{A}\|}{\|\boldsymbol{A}\|} \frac{1}{1 - \|\boldsymbol{A}^{-1}\| \, \|\Delta\boldsymbol{A}\|}$$

Théorème 5 *Pour toute matrice $\boldsymbol{A}$*

1. $cond\{\boldsymbol{A}\} \geq 1$, $cond\{\boldsymbol{A}\} = cond\left\{\boldsymbol{A}^{-1}\right\}$, $cond\{\alpha\boldsymbol{A}\} = cond\{\boldsymbol{A}\}$ *si* $\alpha \neq 0$.

2. $cond_2(\boldsymbol{A}) = \frac{\mu_n(\boldsymbol{A})}{\mu_1(\boldsymbol{A})}$ *où $\mu_1(\boldsymbol{A})$ et $\mu_n(\boldsymbol{A})$ sont respectivement la plus petite et la plus grande valeur singulière de la matrice $\boldsymbol{A}$.*

3. *Si $\boldsymbol{A}$ est une matrice normale; alors*

$$cond_2(\boldsymbol{A}) = \frac{\max_i |\lambda_i(\boldsymbol{A})|}{\min_i |\lambda_i(\boldsymbol{A})|}$$

   où $\lambda_i(\boldsymbol{A})$ sont les valeurs propres de $\boldsymbol{A}$.

4. *Si $\boldsymbol{A}$ est une matrice unitaire ou orthogonale $cond_2(\boldsymbol{A}) = 1$.*

5. *Le conditionnement $cond_2(\boldsymbol{A})$ est invariante par transformation unitaire:*

$$\boldsymbol{UU}^* = I \longrightarrow cond_2(\boldsymbol{AU}) = cond_2(\boldsymbol{UA}) = cond_2(\boldsymbol{U}^*\boldsymbol{AU}) = cond_2(\boldsymbol{A})$$

Dans l'exemple (1) on a

$$\lambda_1 = 0.01015 < \lambda 2 = 0.8431 < \lambda 3 = 3.858 < \lambda 4 = 30.2887$$

$$\mathrm{cond}_2(\boldsymbol{A}) = \frac{\lambda_4}{\lambda_1} = 2984$$

On note que le calcul du $\mathrm{cond}_2(\boldsymbol{A})$ repose sur la connaissance des valeurs propres extrêmes de la matrice $\boldsymbol{A}$. En l'absence de cette information on peut utiliser la relation suivante

$$\|\boldsymbol{A}\|_2 \leq \|\boldsymbol{A}\|_E \leq \sqrt{n}\|\boldsymbol{A}\|_2$$

à condition toutefois qu'on connaisse la matrice $\boldsymbol{A}^{-1}$. Un système linéaire est autant mieux conditionné que le nombre $\mathrm{cond}\,\{\boldsymbol{A}\}$ est voisin de 1 (sous-entendu relativement à une norme matricielle subordonnée particulière). Un système linéaire $\boldsymbol{Ax} = \boldsymbol{b}$ à matrice orthogonale (ou unitaire) est très bien conditionné puisque $\mathrm{cond}_2(\boldsymbol{A}) = 1$. Les transformations orthogonales ou unitaires préservent le conditionnement $\mathrm{cond}_2(\boldsymbol{A})$.

E.8.2 Conditionnement d'un problème de calcul des valeurs singulières d'une matrice

Considérons la matrice carrée d'ordre n

$$\boldsymbol{A}(\epsilon) = \begin{pmatrix} 0 & & & & \epsilon \\ 1 & 0 & & & \\ & 1 & & & \\ & & & . & . \\ & & & 1 & 0 \end{pmatrix}$$

Pour $n = 40$ les valeurs propres de $\boldsymbol{A}(0)$ sont égales à 0 : par contre les valeurs propres de $\boldsymbol{A}(10^{-4})$ ont toutes les mêmes module 10^{-1} (les racine n-ème de ϵ). On constate ainsi que l'on a besoin dans ce problème de calcul des valeurs propres; de mesurer le conditionnement de la matrice.

Théorème 6 (de Bauer-Fike) *Soit $\boldsymbol{A}$ une matrice diagonalisable; $\boldsymbol{P}$ une matrice telle que*

$$\boldsymbol{P}^{-1}\boldsymbol{AP} = diag\,[\lambda_i]$$

et $\|.\|$ une norme matricielle telle que

$$\|diag\,[d_i]\,\| = \max_i \|d_i\|$$

pour toute matrice diagonale. Alors pour toute matrice $\Delta\boldsymbol{A}$

$$sp\,[\boldsymbol{A} + \Delta\boldsymbol{A}] \in \prod_{i=1}^{N} x_i \quad où \quad x_i = \{z \in C \| z - \lambda_i| \leq cond\,\{\boldsymbol{P}\}\,\|\Delta\boldsymbol{A}\|\}$$

E.9 Méthodes récursives de résolution des systèmes d'équations linéaires

Principe

On considère le système d'équations linéaires

$$\boldsymbol{R}\boldsymbol{a} = -\boldsymbol{d}$$

où $\boldsymbol{R}$ est une matrice de dimensions $[N, N]$; $\boldsymbol{a}$ un vecteur de dimension $[N]$ contenant les paramètres inconnus; et $\boldsymbol{d}$ un vecteur de données de dimension $[N]$. Supposons que $\boldsymbol{R}$ soit partitionnée de façon suivante:

$$\boldsymbol{R}_{m+1} = \begin{pmatrix} \boldsymbol{R}_m & \boldsymbol{x}_m \\ \boldsymbol{z}_m^t & y_m \end{pmatrix}, \quad 1 \leq m \leq N-1$$

$$\boldsymbol{R}_m\boldsymbol{a}_m = -\boldsymbol{d}_m, \quad \boldsymbol{R}_{m+1}\boldsymbol{a}_{m+1} = -\boldsymbol{d}_{m+1}, \quad \boldsymbol{d}_{m+1}^t = [\boldsymbol{d}_m^t, d_{m+1}]$$

Appliquant le lemme d'inversion de matrices en bloc à la matrice ainsi partionnée; on obtient:

$$\alpha_m \boldsymbol{R}_{m+1}^{-1} = \begin{pmatrix} \alpha_m \boldsymbol{R}_m^{-1} + \boldsymbol{w}_m \boldsymbol{v}_m^t & \boldsymbol{w}_m \\ \boldsymbol{v}_m^t & 1 \end{pmatrix}, \quad \leq m \leq N-1$$

où

$$\alpha_m = y_m + \boldsymbol{z}_m^t \boldsymbol{w}_m = y_m + \boldsymbol{v}_m^t \boldsymbol{x}_m = \frac{\det[\boldsymbol{R}_{m+1}]}{\det[\boldsymbol{R}_m]}$$

$$\boldsymbol{w}_m = -\boldsymbol{R}_m^{-1}\boldsymbol{x}_m, \quad \boldsymbol{v}_m = -\boldsymbol{R}_m^{-t}\boldsymbol{z}_m$$

Si on s'intéresse à la résolution de l'équation (1.1); plutôt qu'à l'inversion de la matrice (1.2); ces équations peuvent être écrites sous une autre forme:

$$\boldsymbol{a}_{m+1} = \begin{pmatrix} \boldsymbol{a}_m \\ 0 \end{pmatrix} + k_{m+1} \begin{pmatrix} \boldsymbol{w}_m \\ 1 \end{pmatrix}$$

$$k_{m+1} = -\frac{\beta_m}{\alpha_m}, \quad \beta_m = \boldsymbol{v}_m^t \boldsymbol{d}_m + d_{m+1} = \boldsymbol{z}_m^t \boldsymbol{a}_m + d_{m+1}$$

Le coût de calcul de cet algorithme est $O(N^3)$. Son intérêt particulier est dans le fait qu'on dispose de la valeur du déterminant de la matrice à chaque itération; ce qui est utile dans le contrôle de stabilité de l'algorithme pour l'inversion des matrices mal-conditionnées. D'autre part; parmi tous les paramètres utilisés dans cet algorithme, c'est à dire $\boldsymbol{w}_m, \boldsymbol{v}_m, \alpha_m$ et β_m, seul ce dernier dépend des données.

Cas des matrices Symétriques

Dans ce cas on a

$$\boldsymbol{x}_m = \boldsymbol{z}_m = \boldsymbol{r}_m, \quad \text{et} \quad \boldsymbol{w}_m = \boldsymbol{v}_m = \boldsymbol{R}_m^{-1}\boldsymbol{r}_m$$

$$\alpha_m = y_m + \boldsymbol{r}_m^t \boldsymbol{w}_m$$

$$\alpha_m \boldsymbol{R}_{m+1}^{-1} = \begin{pmatrix} \alpha_m \boldsymbol{R}_m^{-1} + \boldsymbol{v}_m \boldsymbol{v}_m^t & \boldsymbol{v}_m \\ \boldsymbol{v}_m^t & 1 \end{pmatrix}, \quad \leq m \leq N-1$$

Cas des matrices MID

$$\boldsymbol{w}_m = -\boldsymbol{R}_m^{-1}\boldsymbol{x}_m, \quad \boldsymbol{w}_{m+1} = -\boldsymbol{R}_{m+1}^{-1}\boldsymbol{x}_{m+1}$$

En utilisant les Matrices MID; c'est à dire:

$$\boldsymbol{x}_{m+1}^t = [\boldsymbol{x}_m^t; x_{m+1}], \quad \boldsymbol{z}_{m+1}^t = [\boldsymbol{z}_m^t; z_{m+1}]; \quad \boldsymbol{y}_{m+1}^t = [\boldsymbol{y}_m^t; y_{m+1}]$$

on obtient une équation de récurrence pour $\boldsymbol{w}_{m+1}$ et $\boldsymbol{v}_{m+1}$ en fonction de $\boldsymbol{w}_m$ et $\boldsymbol{v}_m$.

$$\boldsymbol{w}_{m+1} = \begin{pmatrix} \boldsymbol{w}_m \\ 0 \end{pmatrix} + k_{m+1}^+ \begin{pmatrix} \boldsymbol{w}_m \\ 1 \end{pmatrix}, \quad \boldsymbol{v}_{m+1} = \begin{pmatrix} \boldsymbol{v}_m \\ 0 \end{pmatrix} + k_{m+1}^- \begin{pmatrix} \boldsymbol{v}_m \\ 1 \end{pmatrix}$$

$$k_{m+1}^+ = -\frac{\beta_m^+}{\alpha_m} = \frac{\boldsymbol{v}_m^t \boldsymbol{x}_m + x_{m+1}}{\alpha_m}, \quad k_{m+1}^- = -\frac{\beta_m^-}{\alpha_m} = \frac{\boldsymbol{z}_m^t \boldsymbol{w}_m + z_{m+1}}{\alpha_m}$$

$$\alpha_m = \boldsymbol{z}_m^t \boldsymbol{w}_m + y_{m+1} = \boldsymbol{v}_m^t \boldsymbol{x}_m + y_{m+1}$$

On peut encore simplifier ces équations en trouvant une relation de récurrence pour $\alpha_m, \beta_m^+, \beta_m^-$ et β_m :

$$\alpha_m = \alpha_{m-1} + k_m^+ \beta_{m-1}^- + (y_{m+1} - y_m)$$

$$\beta_m^+ = (x_{m+1} - y_{m+1}) + \alpha_m, \quad \beta_m^- = (z_{m+1} - y_{m+1}) + \alpha_m$$

$$\beta_m = \beta_{m-1}(1 - k_m^-) + (d_{m+1} - d_m)$$

Ces équations simplifient plus si $y_m = x_m$ et / ou $y_m = z_m$. L'inversion de ces matrice nécessitent $O(p^2)$ opérations élémentaires.

Cas des matrices MIP

Il n'existe pas encore des méthodes rapides directes pour inversion des matrices MIP dans le cas général car ces matrices peuvent avoir un rang de déplacement entre $0, \cdots, N$; $[N; N]$ étant les dimensions de la matrice. Par contre un cas particulier de ces matrices est la matrice de Toeplitz $(y_i = y_0, i = 1, .., N)$ non-symétrique $(x_i \neq z_i)$ ou symétrique $(x_i = z_i)$.

Cas des matrices de Toeplitz

Une matrice de Toeplitz satisfait la propriété suivant:

$$\boldsymbol{J}\boldsymbol{R}_m\boldsymbol{J} = \boldsymbol{R}_m^t, \quad m = 1, \cdots, N$$

où $\boldsymbol{J}$ est une matrice d'inversion d'ordre de dimensions $[M, M]$; ce qui vaut dire qu'une matrice de Toeplitz est une matrice symétrique par rapport à sa deuxième diagonale. Si on note par

$$\widetilde{\boldsymbol{x}}_m = [x_1; x_2, \cdots, x_m]^t, \quad \widetilde{\boldsymbol{z}}_m = [x_1; x_2, \cdots, x_m]^t$$

$$\boldsymbol{w}_{m+1} = -\boldsymbol{R}_{m+1}^{-1} \boldsymbol{J} \widetilde{\boldsymbol{x}}_{m+1}, \quad \boldsymbol{w}_{m+1} = -\boldsymbol{J} \boldsymbol{R}_{m+1}^{-t} \widetilde{\boldsymbol{x}}_{m+1}$$

$$\boldsymbol{v}_{m+1}^t = -\widetilde{\boldsymbol{z}}_{m+1}^t \boldsymbol{J} \boldsymbol{R}_{m+1}^{-1} = -\widetilde{\boldsymbol{z}}_{m+1}^t \boldsymbol{R}_{m+1}^{-t} \boldsymbol{J}$$

$$\boldsymbol{J}\boldsymbol{w}_{m+1} = \begin{pmatrix} \boldsymbol{J}\boldsymbol{w}_m \\ 0 \end{pmatrix} + k_{m+1}^+ \begin{pmatrix} \boldsymbol{v}_m \\ 1 \end{pmatrix}, \quad \boldsymbol{J}\boldsymbol{v}_{m+1} = \begin{pmatrix} \boldsymbol{J}\boldsymbol{v}_m \\ 0 \end{pmatrix} + k_{m+1}^- \begin{pmatrix} \boldsymbol{w}_m \\ 1 \end{pmatrix}$$

$$k_{m+1}^+ = -\frac{\beta_m^+}{\alpha_m} = -\frac{\boldsymbol{w}_m^t \widetilde{\boldsymbol{x}}_m + x_{m+1}}{\alpha_m}, \quad k_{m+1}^- = -\frac{\beta_m^-}{\alpha_m} = -\frac{\widetilde{\boldsymbol{z}}_m^t \boldsymbol{w}_m + z_{m+1}}{\alpha_m}$$

$$\alpha_m = \boldsymbol{z}_m^t \boldsymbol{w}_m + y_0 = \boldsymbol{x}_m^t \boldsymbol{v}_m + y_0$$

Notons que les récusions pour $\boldsymbol{w}$ et $\boldsymbol{v}$ sont couplées (ce qui n'était pas le cas pour les matrice MID). Le terme α_m peut être calculé d'une façon plus efficace par

$$\alpha_m = \alpha_{m-1}(1 - k_m^+ k_m^-)$$

On peut obtenir ces mêmes relations pour les matrices de Hankel en remarquant qu'une matrice de Hankel $\boldsymbol{H}$ et une matrice de Toeplitz $\boldsymbol{T}$ sont reliées par $\boldsymbol{H} = \boldsymbol{T}\boldsymbol{J}$ ou $\boldsymbol{H} = \boldsymbol{J}\boldsymbol{T}$. Ainsi les trois relations suivantes décrivent un même système d'équations linéaires:

$$\boldsymbol{H}\boldsymbol{a} = -\boldsymbol{d}, \quad \boldsymbol{T}(\boldsymbol{J}\boldsymbol{a}) = -\boldsymbol{d}, \quad \boldsymbol{T}\boldsymbol{a} = -\boldsymbol{J}\boldsymbol{d}$$

il suffit alors de remplacer $\boldsymbol{d}$ par $\boldsymbol{J}\boldsymbol{d}$ pour obtenir les relations nécessaires pour des matrices de Hankel.

Cas des matrices de Toeplitz Symétriques

Algorithme de Levinson :
Dans ce cas on a la relation $\boldsymbol{w} = \boldsymbol{v}$ ce qui va permettre de simplifier encore plus les relations précédentes. On a

$$k^+ = k^- = k$$

$$\boldsymbol{J}\boldsymbol{w}_{m+1} = \begin{pmatrix} \boldsymbol{J}\boldsymbol{w}_m \\ 0 \end{pmatrix} + k_{m+1} \begin{pmatrix} \boldsymbol{v}_m \\ 1 \end{pmatrix}$$

$$\beta = \boldsymbol{v}_{m+1}^t \widetilde{\boldsymbol{x}}_m + x_{m+1}, \quad \alpha_m = \alpha_{m-1}(1 - k_m^2)$$

Algorithme de Durbin :
Dans le cas particulier où le le vecteur $\boldsymbol{d}$peut être écrit sous la forme

$$\boldsymbol{d}_N^t = [\boldsymbol{d}_{N-1}^t + d_N]$$

comme c'est le cas du problème de la modélisation AR; on a la relation

$$\boldsymbol{w}_m = \boldsymbol{J}\boldsymbol{a}_m, \quad m = 1, 2, \cdots, N-1$$

Cas des matrices circulantes

Dans le cas des matrices circulantes on a

$$\boldsymbol{z}_{N-1} = \boldsymbol{J}\boldsymbol{x}_{N-1}, \quad \boldsymbol{v}_{N-1} = \boldsymbol{J}\boldsymbol{w}_{N-1}$$

Il faut noter que l'on peut diagonaliser les matrices circulantes par TFD et obtenir des algorithmes non récursifs.

Cas des matrices de Toeplitz en bande

E.10 Diverses propriétés

Lemme de factorisation d'une matrice

Soit $\boldsymbol{A}$ une matrice de dimension $[N, P]$. On peut factoriser la matrice $[\boldsymbol{I}_N \pm \boldsymbol{A}\boldsymbol{A}^t]$ de façon suivante:

$$\boldsymbol{I}_N \pm \boldsymbol{A}\boldsymbol{A}^t = [\boldsymbol{I}_N \pm \boldsymbol{A}\boldsymbol{R}^{-1}\boldsymbol{A}^t].[\boldsymbol{I}_N \pm \boldsymbol{A}\boldsymbol{R}^{-1}\boldsymbol{A}^t]^t$$

où $\boldsymbol{R} = \boldsymbol{I}_{\boldsymbol{P}} + [\boldsymbol{I}_{\boldsymbol{P}} \pm \boldsymbol{A}^t\boldsymbol{A}]^{1/2}$.

Racines carrées d'une matrice définie positive

Toute matrice définie positive admet une racine carrée. On appelle racine carrée d'une matrice $\boldsymbol{A}$ toute matrice carrée $\boldsymbol{S}$, telle que

$$\boldsymbol{A} = \boldsymbol{S}\boldsymbol{S}^t, \quad \boldsymbol{S} = \boldsymbol{A}^{1/2}, \quad \boldsymbol{S}^t = \boldsymbol{A}^{t/2}.$$

Une telle factorisation n'est pas unique. Si $\boldsymbol{T}$ est orthogonale ($\boldsymbol{T}\boldsymbol{T}^t = \boldsymbol{I}$), alors $\boldsymbol{S}\boldsymbol{T}$ est aussi racine carrée de $\boldsymbol{A}$ puisque $\boldsymbol{A} = \boldsymbol{S}\boldsymbol{T}\boldsymbol{T}^t\boldsymbol{S}^t = \boldsymbol{S}\boldsymbol{S}^t$.

Décomposition de Cholesky d'une matrice définie positive

Soit $\boldsymbol{A}$ une matrice symétrique, réelle, et définie positive d'ordre N. Il existe alors une matrice $\boldsymbol{O}$ orthogonale réelle d'ordre N telle que $\boldsymbol{D} = \boldsymbol{O}^{-1}\boldsymbol{A}\boldsymbol{O} = \boldsymbol{O}^t\boldsymbol{A}\boldsymbol{O}$ soit diagonale. De plus, si $\boldsymbol{A}$ est de rang $K < N$, elle admet une décomposition de la forme

$$\boldsymbol{A} = \boldsymbol{S}_1\boldsymbol{S}_1^t - \boldsymbol{S}_2\boldsymbol{S}_2^t$$

où $\boldsymbol{S}_1$ et $\boldsymbol{S}_2$ sont de dimensions respectives $[N, \alpha]$ et $[N, K - \alpha]$, α étant le nombre de valeurs propres positives non nulles de $\boldsymbol{A}$. La décomposition de Cholesky consiste à imposer aux racines carrées d'être des matrices triangulaires. La décomposition est alors unique.

Le rang de déplacement d'une matrice

Le rang de déplacement d'une matrice est le rang de la différence entre la matrice initiale et la matrice décalée d'un cran le long de sa diagonale principale. On définit deux rangs de déplacement, inférieur δ_- et supérieur δ_+, à partir des deux déplacements possibles (vers le haut-gauche, et vers le bas-droite):

$$\delta_+(\boldsymbol{R}) = \text{rang}\left[\boldsymbol{R} - \boldsymbol{Z}\boldsymbol{R}\boldsymbol{Z}^t\right], \quad \delta_-(\boldsymbol{R}) = \text{rang}\left[\boldsymbol{R} - \boldsymbol{Z}^t\boldsymbol{R}\boldsymbol{Z}\right]$$

où $\boldsymbol{Z}$ est une matrice de décalage vers le bas donnée par:

$$\boldsymbol{Z} = \begin{pmatrix} 0 & . & . & . & 0 \\ 1 & 0 & . & . & 0 \\ 0 & 1 & 0 & . & 0 \\ . & 0 & 1 & 0 & 0 \\ . & . & 0 & 1 & 0 \end{pmatrix}$$

Pour une matrice quelconque $\boldsymbol{R}$ on peut toujours construire deux partitions :

$$\boldsymbol{R} = \begin{pmatrix} r_0 & \boldsymbol{x}^t \\ \boldsymbol{y} & \boldsymbol{R}_i \end{pmatrix} = \begin{pmatrix} \boldsymbol{R}_s & \boldsymbol{u} \\ \boldsymbol{v}^t & s_0 \end{pmatrix}$$

Ainsi, pour cette matrice on a :

$$\boldsymbol{R} - \boldsymbol{Z}\boldsymbol{R}\boldsymbol{Z}^t = \begin{pmatrix} \boldsymbol{R}_i - \boldsymbol{R}_s & \boldsymbol{u} \\ \boldsymbol{v}^t & s_0 \end{pmatrix}, \quad \text{et} \quad \boldsymbol{R} - \boldsymbol{Z}^t\boldsymbol{R}\boldsymbol{Z} = \begin{pmatrix} r_0 & \boldsymbol{x}^t \\ \boldsymbol{y} & \boldsymbol{R}_i - \boldsymbol{R}_s \end{pmatrix}$$

Dans le cas des matrices symétriques on a $\boldsymbol{u} = \boldsymbol{v} = \boldsymbol{s}$, $\boldsymbol{x} = \boldsymbol{y} = \boldsymbol{r}$. On peut alors construire deux partitions de la matrice initiale $\boldsymbol{R}$:

$$\boldsymbol{R} = \begin{pmatrix} r_0 & \boldsymbol{r}^t \\ \boldsymbol{r} & \boldsymbol{R}_i \end{pmatrix} = \begin{pmatrix} \boldsymbol{R}_s & \boldsymbol{s} \\ \boldsymbol{s}^t & s_0 \end{pmatrix}$$

Dans ce cas on a

$$\boldsymbol{R} - \boldsymbol{Z}\boldsymbol{R}\boldsymbol{Z}^t = \begin{pmatrix} \boldsymbol{R}_i - \boldsymbol{R}_s & \boldsymbol{s} \\ \boldsymbol{s}^t & s_0 \end{pmatrix} \quad \text{et} \quad \boldsymbol{R} - \boldsymbol{Z}^t\boldsymbol{R}\boldsymbol{Z} = \begin{pmatrix} r_0 & \boldsymbol{r}^t \\ \boldsymbol{r} & \boldsymbol{R}_i - \boldsymbol{R}_s \end{pmatrix}$$

Pour une matrice de Toeplitz, on vérifie facilement, qu'on a: $\delta_+(\boldsymbol{T}) = \delta_-(\boldsymbol{T}) = 2$ On démontre aussi facilement [] que :

$$\delta_+(\boldsymbol{R}) = \delta_- \boldsymbol{R}^{-1} \quad \text{et} \quad \delta_-(\boldsymbol{R}) = \delta_+ \boldsymbol{R}^{-1} \det\left[\delta_+(\boldsymbol{R}) - \delta_-(\boldsymbol{R})\right] \leq 2$$

Distance à Toeplitz

Le rang de déplacement $\delta_-(\boldsymbol{R})$ ou $\delta_+(\boldsymbol{R})$ ne constitue pas à proprement parler une mesure de la distance à Toeplitz puisque $\delta_-(\boldsymbol{T}) = \delta_-(\boldsymbol{T}) = 2$. C'est la raison pour laquelle on utilise aussi une définition différente du rang de déplacement qui fourni directement une *distance à Toeplitz*.

Soit $\boldsymbol{R}$ une matrice-bloc quelconque de dimensions $[NP, NP]$ $\boldsymbol{R} = [\boldsymbol{R}_{ji}]$, $\quad i, j = 1, \cdots, N$ où $\boldsymbol{R}_{ij}$ sont des matrices de dimensions $[P, P]$. Le déplacement est alors défini par

$$\delta\boldsymbol{R} = [\boldsymbol{R}_{(i+1),(j+1)} - \boldsymbol{R}_{ij}], \quad i, j = 1, \cdots, N-1$$

et le rang de déplacement devient $\alpha =$ partie entière de $[\frac{\text{rang}[\delta\boldsymbol{R}]}{P}]$.

- Quand les éléments de $\boldsymbol{R}$ sont scalaires, $P = 1$ et $\alpha = \text{rang}\,[\delta\boldsymbol{R}]$.
- Si $\boldsymbol{R}$ est une matrice quelconque, $\delta\boldsymbol{R}$ peut être de rang plein et alors $\alpha = N - 1$.
- Si $\boldsymbol{R}$ est Toeplitz, alors $\alpha = 0$. On a en général $0 \leq \alpha \leq N - 1$, et α peut servir directement de mesure de la distance de $\boldsymbol{R}$ à Toeplitz.