

Calcul scientifique et Analyse de données

Introduction

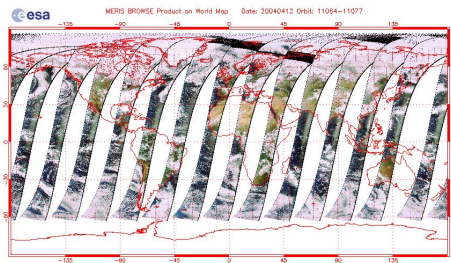
2018–2019

Première partie I

Motivations

Exemple 1 : Reconstruction de données manquantes

- Missions satellitaires relatives à l'observation de l'environnement.



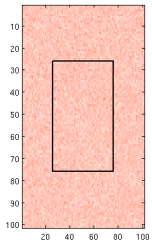
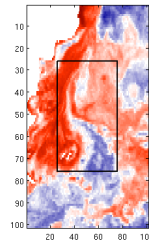
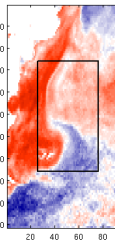
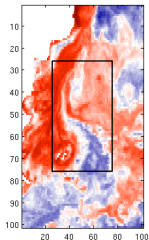
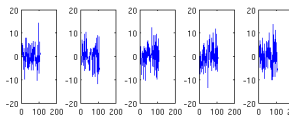
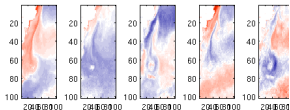
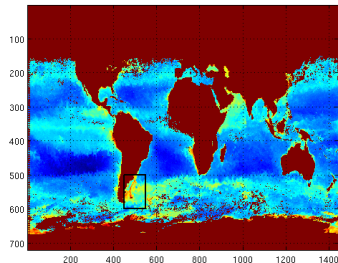
Cycle journalier de MERIS (Envisat) en avril 2004 (ESA)

- Instruments radiométriques (couleur de l'eau) dans le domaine du visible/quasi-infrarouge
 - ▷ "Trous" dans les données causés par la nuit, la couverture nuageuse.
- Peut-on remplir les trous ? Comment ?

Exemple 1 : Reconstruction de données manquantes

Modélisation du problème

- Collection d'images satellite : tableau de données.
- Réduction de dimension du problème : construction de sous-espaces "pertinents".
 - ▷ Un problème de valeurs propres.
 - ▷ Analyse en composantes principales, décomposition SVD, méthode de la puissance itérée.
- Projection de la nouvelle image sur ces sous-espaces.

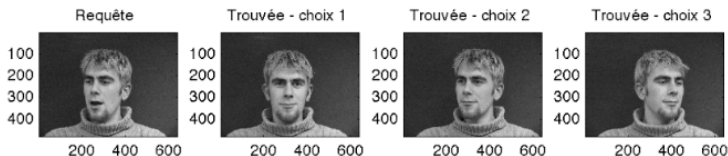
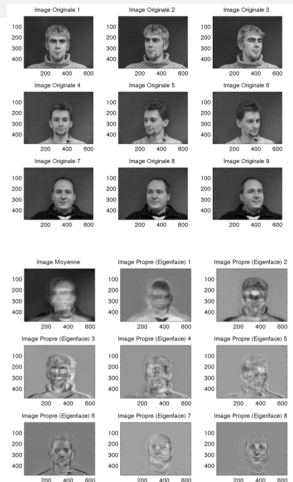


Exemple 2 : Reconnaissance de visages

Modélisation du problème

- Collection d'images de visages.
- Réduction de dimension du problème : construction de sous-espaces "pertinents".
 - ▷ Un problème de valeurs propres.
 - ▷ Analyse en composantes principales, décomposition SVD, méthode de la puissance itérée.
- Classification de la nouvelle image.
- **Projet 2017-2018**

Collections d'images
Données "comprimées"



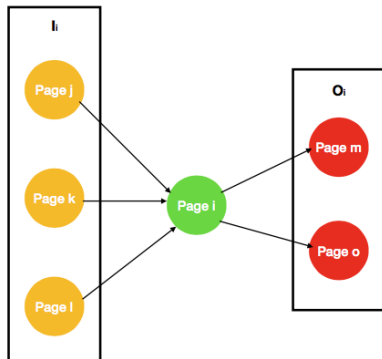
Exemple 3 : Moteurs de recherche

Classement des pages Web (*page ranking*)

- Internet : graphe orienté.
- Importance relative d'une page dans le graphe de l'Internet :
 - ▷ I_i l'ensemble des pages Web ayant un lien vers la page i .
 - ▷ O_i l'ensemble des pages Web vers lesquelles la page i a un lien. $N_i = |O_i|$.
 - ▷ $r \in \mathbb{R}^n$ le vecteur des importances relatives des pages Web défini par $r = Qr$ avec

$$Q_{ij} = \begin{cases} 1/N_j & \text{si } j \in I_i \\ 0 & \text{sinon} \end{cases}$$

- Un problème de recherche de vecteurs propres.
 - ▷ Méthode de la puissance itérée.

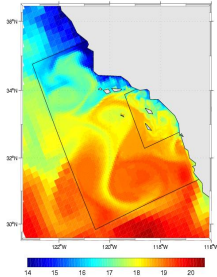
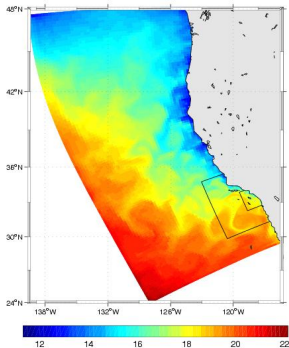


Exemple 4 : Prédiction de l'évolution de notre milieu naturel

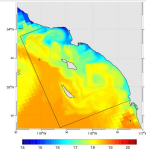
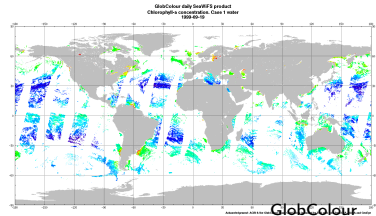
Assimilation de données

Ensemble des techniques permettant de combiner l'*information mathématique* contenue dans les équations et l'*information physique* provenant des observations, en vue de *reconstituer au mieux l'état d'un système*.

Modèles numériques



Observations



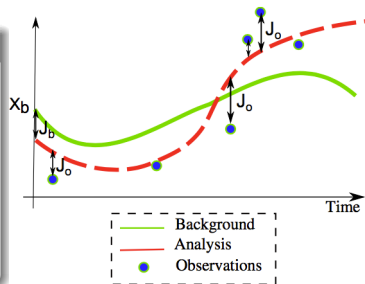
ROMS-AGRIF (UCLA/IRD, Rutgers), Penven, Marchesiello, McWilliams, Debreu.

Exemple 4 : Prédiction de l'évolution de notre milieu naturel

Minimisation d'une fonctionnelle

- Fonction de coût 4D-Var :

$$J(\mathbf{x}_0) = \underbrace{\frac{1}{2}(\mathbf{x}_0 - \mathbf{x}^b)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}^b)}_{\text{écart à l'ébauche}} + \underbrace{\frac{1}{2} \sum_{i=0}^N (H_i[M_{0 \rightarrow i}(\mathbf{x}_0)] - \mathbf{y}_i)^T \mathbf{R}_i^{-1} (H_i[M_{0 \rightarrow i}(\mathbf{x}_0)] - \mathbf{y}_i)}_{\text{écart aux observations}}$$



Résolution en pratique

- Résolution de systèmes linéaires.
 - ▷ Résolution d'équations aux dérivées partielles (modèles numériques).
 - ▷ Minimisation par méthodes de gradients (Gauss-Newton, etc..).
- Moindres carrés, factorisations LU, Cholesky, steepest descent, méthodes des gradients conjugués.
- Réduction des dimensions des problèmes.
 - ▷ Décomposition en valeurs singulières, factorisation QR.

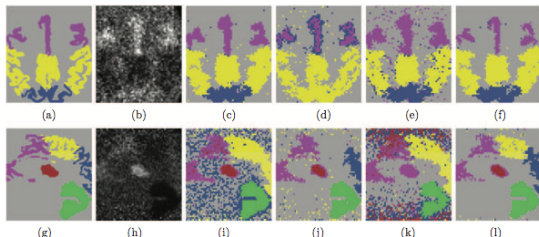
Exemple 5 : Partitionnement de données

- Division d'un ensemble de données en différents groupes "homogènes".
 - ▷ Etudes de marchés, reconnaissance de communautés, écologie, segmentation d'images, etc..
- Apprentissage non supervisé, méthodes de K-means, algorithme EM.

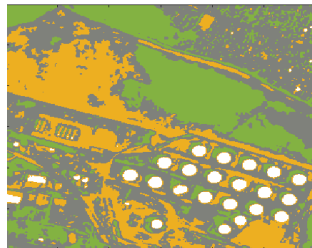
Segmentation d'images (CNES).



Mouysset et al. (2013)



Partitionnement de données médicales (cerveau).



Deuxième partie II

Organisation de l'UE

Organisation de l'UE

Calcul scientifique (CS)

- Algèbre linéaire
- Résolution numérique de systèmes linéaires.
- Calcul de valeurs et vecteurs propres
- Erreur numérique
- Examen : 25% de la note de l'UE

Analyse de données (AD)

- Réduction des dimensions : pouvoir exploiter les données.
- Moindres carrées.
- Apprentissage supervisé.
- Apprentissage non supervisé.
- Examen : 25% de la note de l'UE.

Projet CS et AD

- Projet commun aux deux cours.
 - ▶ Modélisation d'un problème d'analyse de données et développement d'algorithmes pour sa résolution.
- Groupes de 3/4 étudiants.
- Trois séances encadrées : travaux en autonomie entre les séances.
- 50% de la note de l'UE.

Calendrier indicatif

Mois	Calcul Scientifique	Analyse de Données	Projet
Janvier	CM-1 :2 : SVD et applications CM-3 : SVD ; CTD-1 : valeurs propres	CTD-1 : ACP TP-1 : ACP	
Février	CM-4 : Erreurs TP-1 : Erreurs CTD-2 :3 : Systèmes linéaires	CTD-2 : MC TP-2 : MC ; CM-1 : Classif. sup. CM-2, TP-3 : Classif. sup.	TP-1
Mars	CTD-4 : Systèmes linéaires Vacances	Vacances	TP-2
Avril	TP-2 : Systèmes linéaires CTD-5 : Facto QR CTD-6 : Facto QR Vacances Vacances	CM-3 : Classif. non sup. CM-4 : Classif. non sup. TD-1 TP-4 : Classif. non sup. Vacances Vacances	TP-3
Mai	TD-1 TD-2 Examen CM-5 : Ouvertures	Examen	

Troisième partie III

Rappels d'algèbre linéaire

Rappel sur les matrices : $A \in \mathcal{M}_n(\mathbb{R})$

- $A = [a_{ij}]$ $\begin{cases} i = 1, \dots, n & \text{indices de lignes} \\ j = 1, \dots, n & \text{indices de colonnes} \end{cases}$
- Matrice transposée ${}^tA = A^T = [a_{ji}]$
 $(A \cdot B)^T = B^T \cdot A^T \quad (A \cdot B)^{-1} = B^{-1} \cdot A^{-1} \quad (A^{-1})^T = (A^T)^{-1}$
- $(\cdot|\cdot)$ produit scalaire Euclidien : $(x|A \cdot y) = (A^T \cdot x|y) = x^T \cdot A \cdot y$
- A symétrique : $A = A^T$ ($a_{ij} = a_{ji} \quad i, j = 1, \dots, n$)
- Déterminant de A : $\det(A) = |A|$
 $\forall \alpha \in \mathbb{R} \quad |\alpha A| = \alpha^n |A| \quad |A| = |A^T|$
 $|A \cdot B| = |A| \cdot |B| \quad |A^{-1}| = 1/|A|$

Rappel sur les matrices : $A \in \mathcal{M}_n(\mathbb{R})$

$$\begin{aligned} A \text{ inversible} &\iff |A| \neq 0 \\ &\iff \text{rang}(A) = n \\ &\iff A \text{ régulière} \end{aligned}$$

$$\begin{aligned} A \text{ définie positive} &\iff \forall x \in \mathbb{R}^{n*} \quad x^T \cdot A \cdot x = (x|A \cdot x) > 0 \\ &\iff \forall x \in \mathbb{R}^n \quad (x|A \cdot x) \geq 0 \\ &\quad \text{et } (x|A \cdot x) = 0 \implies x = 0 \end{aligned}$$

$$A \text{ orthogonale} \iff A^{-1} = A^T$$

$$\begin{aligned} \lambda \in \mathbb{C} \text{ valeur propre de } A &\iff \exists x \in \mathbb{C}^{n*}, \text{ appelé } \textit{vecteur propre} \\ &\quad \text{t.q. } A \cdot x = \lambda x \end{aligned}$$

Opérations élémentaires sur les matrices

Les matrices élémentaires dans $\mathcal{M}_n(\mathbb{K})$

- Base canonique de $\mathcal{M}_{m,n}(\mathbb{K})$: $E_{i,j} = [\delta_{p,i}\delta_{q,j}]_{1 \leq p \leq m, 1 \leq q \leq n}$
- **Matrice de permutation** $P_\sigma = \sum_{j=1}^n E_{\sigma(j),j}$, avec σ une permutation de $\{1, \dots, n\}$.
On a $P_\sigma^{-1} = P_{\sigma^{-1}}$.
- **Matrice de transvection** $T_{i,j}(\lambda) = I_n + \lambda E_{i,j}$, avec $i \neq j$ et $\lambda \neq 0$
On a $T_{i,j}(\lambda)^{-1} = T_{i,j}(-\lambda)$.
- **Matrice de dilatation** $D_i(\lambda) = I_n + (\lambda - 1)E_{i,i}$ avec $\lambda \neq 0$.
On a $D_i(\lambda)^{-1} = D_i(\frac{1}{\lambda})$

Opérations élémentaires sur les matrices

Opérations élémentaires sur les lignes : multiplication à gauche

- $E_{i,j}.A$ est la matrice dont toutes les lignes sont nulles sauf la i -ème qui est la j -ème ligne de A .
- $P_\sigma.A$ est la matrice obtenue en effectuant la permutation σ^{-1} sur les lignes de A .
- $T_{i,j}(\lambda).A : L_i \leftarrow L_i + \lambda L_j$, avec L_i et L_j les i ème et j ème lignes de A .
- $D_i(\lambda).A : L_i \leftarrow \lambda L_i$.

Opérations élémentaires sur les colonnes : multiplication à droite

- $A.E_{i,j}$ est la matrice dont toutes les colonnes sont nulles sauf la j -ème qui est la i -ème colonne de A .
- $A.P_\sigma$ est la matrice obtenue en effectuant la permutation σ sur les colonnes de A .
- $A.T_{i,j}(\lambda) : C_j \leftarrow C_j + \lambda C_i$, avec C_i et C_j les i ème et j ème colonnes de A .
- $A.D_j(\lambda) : C_j \leftarrow \lambda C_j$.

Normes matricielles

$\mathcal{M}_n(\mathbb{R})$ espace vectoriel de dimension finie : toutes les normes sont équivalentes.

- Norme la plus usuelle : **norme matricielle induite par une norme vectorielle** $\|\cdot\|$ de \mathbb{R}^n

$$A \in \mathcal{M}_n(\mathbb{R}) \quad \|A\| = \sup_{\|x\| \neq 0} \frac{\|Ax\|}{\|x\|} = \sup_{\|x\|=1} \|Ax\|$$

$$\textbf{Propriétés : } \|Ax\| \leq \|A\| \cdot \|x\| \quad \|A \cdot B\| \leq \|A\| \cdot \|B\|$$

Norme la plus usuelle : norme matricielle induite par une p -norme vectorielle

$$(\forall x \in \mathbb{R}^n, \|x\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}}).$$

On a $\forall A \in \mathcal{M}_n(\mathbb{R})$,

- $\|A\|_1 = \max_{j=1:n} \sum_{i=1}^n |a_{ij}|$.
- $\|A\|_2 = \sqrt{\rho(A^T A)}$, avec ρ le rayon spectral.

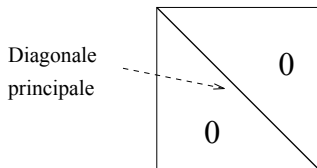
$$\forall x \in \mathbb{R}^n, \|x\|_\infty = \max_{i=1:n} |x_i|, \quad \forall A \in \mathcal{M}_n(\mathbb{R}), \|A\|_\infty = \max_{i=1:n} \sum_{j=1}^n |a_{ij}|.$$

- Autres normes usuelles :

norme Max : $\|A\|_{\max} = \max_{i,j} |a_{i,j}|$ (pas sous-multiplicative !)

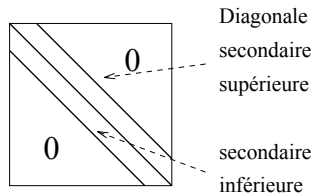
$$\text{norme de Frobenius : } \|A\|_F = \left(\sum_{i=1}^m \sum_{j=1}^n |a_{i,j}|^2 \right)^{\frac{1}{2}} = \sqrt{\text{tr}(A^* A)}$$

Structures par blocs

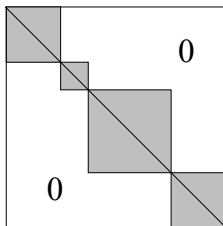


Diagonale
principale

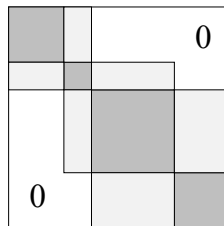
Matrice diagonale



Matrice tridiagonale



Matrice diagonale
par blocs



Matrice tridiagonale
par blocs

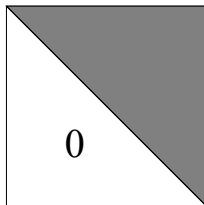
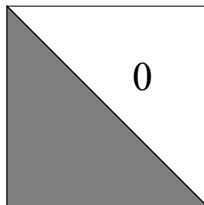
Structures particulières

Matrice triangulaire

Soit A triangulaire inférieure (resp. supérieure) :

- Si A inversible alors A^{-1} est triangulaire inférieure (resp. supérieure).
- $|A| = \text{produit des coefficients diagonaux}$.
- Le produit de 2 matrices triangulaires inférieures est triangulaire inférieur.

Matrice
triangulaire
inférieure



Matrice
triangulaire
supérieure

Structures particulières

Matrice triangularisable

Une matrice A est dite triangularisable si elle est semblable à une matrice triangulaire supérieure.

Prop. :

- $\forall A \in \mathcal{M}_n(\mathbb{K})$, A triangularisable \iff le polynôme caractéristique de A est scindé sur \mathbb{K} .
- $\forall A \in \mathcal{M}_n(\mathbb{C})$ est triangularisable.

Théorème de triangularisation unitaire de Schur

$\forall A \in \mathcal{M}_n(\mathbb{C})$, il existe $U \in \mathcal{M}_n(\mathbb{C})$ unitaire ($U^* = U^{-1}$) telle que U^*AU est triangulaire supérieure avec les valeurs propres de A sur la diagonale.

Remarque :

- Si $A \in \mathcal{M}_n(\mathbb{R})$, avec des valeurs propres toutes réelles, alors $U \in \mathcal{M}_n(\mathbb{R})$ est orthogonale.

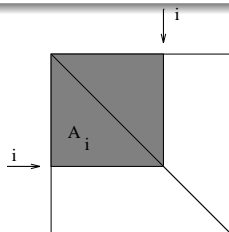
Structures particulières

Matrice symétrique

- Les valeurs propres sont réelles (non nécessairement distinctes)
- Si u_1 est un vecteur propre associé à λ_1 , si u_2 est un vecteur propre associé à λ_2 (avec $\lambda_1 \neq \lambda_2$), alors $(u_1|u_2) = 0$

Matrice symétrique définie positive

- Les valeurs propres sont strictement positives (c'est une CNS pour montrer le caractère défini positif d'une matrice symétrique)
- Le déterminant et les coefficients diagonaux sont strictement positifs (la matrice est donc inversible)
- Les sous-matrices principales sont symétriques définies positives



A sous-matrice
principale de
rang i

Systèmes à structure particulière

Systèmes diagonaux, diagonaux par blocs

- Algorithmique naturelle

Systèmes tridiagonaux, tridiagonaux par blocs

- Simplification d'un algorithme de résolution général
ou
- Méthode spécifique comme les techniques de réduction

Systèmes triangulaires inférieurs

- Algorithme sans reports pour un accès à la matrice par lignes
ou
- Algorithme avec reports pour un accès à la matrice par colonnes

Principes

Le nombre d'opérations ne dépend que de la dimension de la matrice (il est indépendant de la valeur des coefficients). Résultat exact aux erreurs numériques près.

Les plus fréquentes : **méthodes de factorisation**

$$A \cdot x = b \rightarrow A_1 \cdot A_2 \cdot A_3 \cdot x = b ; A_1 \cdot y = b ; A_2 \cdot z = y ; A_3 \cdot x = z$$

- Cas général : factorisation de Gauss
- Matrice symétrique définie positive : factorisation de Cholesky

Contraintes des méthodes de factorisation :

- La factorisation met en évidence des matrices à structure simple (A_1, A_2, A_3 triangulaires ou diagonales, tridiagonales, orthogonales...)
- Une seule structure de données matricielle de mémorisée (les coefficients de A_1, A_2, A_3 se substituent à certains des coefficients de A)

Principe

On cherche une suite $x^{(p)}$ de \mathbb{R}^n convergeant vers la solution de $A \cdot x = b$

$$\forall x^{(0)} \quad x^{(p+1)} = H(x^{(p)}).$$

Propriétés

- La matrice A n'est jamais modifiée.
- Problème du suivi de la convergence et du choix du test d'arrêt.
- La solution obtenue n'est pas exacte.
- La matrice doit vérifier des conditions de convergence.
- La vitesse de convergence dépend de la valeur des coefficients de la matrice.