

Mathématiques assistées par ordinateur

Chapitre 9 : Calcul matriciel et algèbre linéaire

Michael Eisermann

Mat249, DLST L2S4, Année 2008-2009

www-fourier.ujf-grenoble.fr/~eiserm/cours#mao

Document mis à jour le 6 juillet 2009



Sommaire

- 1 Résolution de systèmes d'équations linéaires
- 2 Réduction des endomorphismes
- 3 Méthodes approchées itératives
- 4 Comment fonctionne Google ?

Sommaire

- 1** Résolution de systèmes d'équations linéaires
 - Systèmes d'équations linéaires, l'algorithme de Gauss
 - Calcul matriciel : addition, multiplication, inversion, déterminant
 - Stabilité numérique, conditionnement d'une matrice
- 2 Réduction des endomorphismes
- 3 Méthodes approchées itératives
- 4 Comment fonctionne Google ?

Systemes d'equations lineaires

Dans la suite nous fixons un corps \mathbb{K} (par exemple \mathbb{Q} , \mathbb{R} , ou \mathbb{C}).

Systemes d'equations lineaires

Dans la suite nous fixons un corps \mathbb{K} (par exemple \mathbb{Q} , \mathbb{R} , ou \mathbb{C}).
Nous souhaitons resoudre un *systeme d'equations lineaires* :

$$\left\{ \begin{array}{lcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & y_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = & y_2 \\ & & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = & y_m \end{array} \right.$$

Systèmes d'équations linéaires

Dans la suite nous fixons un corps \mathbb{K} (par exemple \mathbb{Q} , \mathbb{R} , ou \mathbb{C}).
Nous souhaitons résoudre un *système d'équations linéaires* :

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & y_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = & y_2 \\ & & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = & y_m \end{cases}$$

On écrit ce système plus succinctement comme $Ax = y$ où

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}.$$

Matrices triangulaires et échelonnées

On cherche à résoudre un système d'équations linéaires $Ax = y$.

Matrices triangulaires et échelonnées

On cherche à résoudre un système d'équations linéaires $Ax = y$.

Si A est *triangulaire*, la solution est immédiate : il suffit de remonter.

$$A = \begin{pmatrix} 1 & * & * & * \\ 0 & 1 & * & * \\ 0 & 0 & 1 & * \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{voire} \quad A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Matrices triangulaires et échelonnées

On cherche à résoudre un système d'équations linéaires $Ax = y$.

Si A est *triangulaire*, la solution est immédiate : il suffit de remonter.

$$A = \begin{pmatrix} 1 & * & * & * \\ 0 & 1 & * & * \\ 0 & 0 & 1 & * \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{voire} \quad A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Plus généralement la solution est facile si A est *échelonnée* :

$$A = \begin{pmatrix} 1 & * & * & * & * & * & * \\ 0 & 0 & 1 & * & * & * & * \\ 0 & 0 & 0 & 1 & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{voire} \quad A = \begin{pmatrix} 1 & * & 0 & 0 & * & * & 0 \\ 0 & 0 & 1 & 0 & * & * & 0 \\ 0 & 0 & 0 & 1 & * & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Matrices triangulaires et échelonnées

On cherche à résoudre un système d'équations linéaires $Ax = y$.

Si A est *triangulaire*, la solution est immédiate : il suffit de remonter.

$$A = \begin{pmatrix} 1 & * & * & * \\ 0 & 1 & * & * \\ 0 & 0 & 1 & * \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{voire} \quad A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Plus généralement la solution est facile si A est *échelonnée* :

$$A = \begin{pmatrix} 1 & * & * & * & * & * & * \\ 0 & 0 & 1 & * & * & * & * \\ 0 & 0 & 0 & 1 & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{voire} \quad A = \begin{pmatrix} 1 & * & 0 & 0 & * & * & 0 \\ 0 & 0 & 1 & 0 & * & * & 0 \\ 0 & 0 & 0 & 1 & * & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

« *Solution générale = solution particulière + solutions homogènes.* »

Opérations élémentaires

Objectif : mettre sous forme échelonnée une matrice donnée

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} = \begin{pmatrix} L_1 \\ \vdots \\ L_m \end{pmatrix}.$$

Opérations élémentaires

Objectif : mettre sous forme échelonnée une matrice donnée

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} = \begin{pmatrix} L_1 \\ \vdots \\ L_m \end{pmatrix}.$$

Pour ceci on effectue des *opérations élémentaires* sur les lignes :

- $L_i \leftrightarrow L_j$ échanger la ligne i et la ligne j ,
- $L_i \leftarrow \lambda L_i$ multiplier la ligne i par un facteur inversible λ ,
- $L_i \leftarrow L_i + \lambda L_j$ ajouter un multiple de la ligne j à la ligne i .

L'algorithme de Gauss : idée

D'abord on choisit un pivot $a_{j1} \neq 0$ et on le place en tête :

$$\begin{pmatrix} a_{11} & a_{21} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \xrightarrow{L_1 \leftrightarrow L_j} \begin{pmatrix} a_{11} & a_{21} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

Ensuite on annule la première colonne :

$$\begin{array}{c} \xrightarrow[\substack{a_{11} \neq 0 \\ L_1 \leftarrow a_{11}^{-1} L_1}]{} \\ \\ \xrightarrow[\substack{j=2, \dots, m \\ L_j \leftarrow L_j - a_{j1} L_1}]{} \end{array} \begin{pmatrix} 1 & a_{21} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$
$$\begin{pmatrix} 1 & a_{21} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

L'algorithme de Gauss : idée

D'abord on choisit un pivot $a_{j1} \neq 0$ et on le place en tête :

$$\begin{pmatrix} a_{11} & a_{21} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \xrightarrow{L_1 \leftrightarrow L_j} \begin{pmatrix} a_{11} & a_{21} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

Ensuite on annule la première colonne :

$$\begin{array}{c} \xrightarrow[\substack{a_{11} \neq 0 \\ L_1 \leftarrow a_{11}^{-1} L_1}]{} \\ \xrightarrow[\substack{j=2, \dots, m \\ L_j \leftarrow L_j - a_{j1} L_1}]{} \end{array} \begin{pmatrix} 1 & a_{21} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

On itère cette méthode sur la sous-matrice $\{2, \dots, m\} \times \{2, \dots, n\}$.

Avertissement numérique : le bon pivot

Soit à résoudre le système linéaire suivant avec un paramètre $\varepsilon \approx 0$:

$$\begin{cases} \varepsilon x + 1.0y = 1.0 \\ x + 2.0y = 3.0 \end{cases}$$

Avertissement numérique : le bon pivot

Soit à résoudre le système linéaire suivant avec un paramètre $\varepsilon \approx 0$:

$$\begin{cases} \varepsilon x + 1.0y = 1.0 \\ x + 2.0y = 3.0 \end{cases}$$

La solution exacte est $x = \frac{1}{1-2\varepsilon} \approx 1$ et $y = \frac{1-3\varepsilon}{1-2\varepsilon} \approx 1$. (Exercice !)

Avertissement numérique : le bon pivot

Soit à résoudre le système linéaire suivant avec un paramètre $\varepsilon \approx 0$:

$$\begin{cases} \varepsilon x + 1.0y = 1.0 \\ x + 2.0y = 3.0 \end{cases}$$

La solution exacte est $x = \frac{1}{1-2\varepsilon} \approx 1$ et $y = \frac{1-3\varepsilon}{1-2\varepsilon} \approx 1$. (Exercice !)

Supposons d'abord que l'on utilise le pivot 1, comme il se doit :

$$\begin{array}{l} \text{On obtient} \\ \text{puis} \end{array} \begin{cases} x + 2.0y = 3.0 \\ \varepsilon x + 1.0y = 1.0 \end{cases}$$
$$\begin{cases} x + 2.0y = 3.0 \\ (1.0 - 2.0\varepsilon)y = (1.0 - 3.0\varepsilon) \end{cases}$$

Avertissement numérique : le mauvais pivot

On considère toujours le même système linéaire où $\varepsilon \approx 0$.

$$\begin{cases} \varepsilon x + 1.0y = 1.0 \\ x + 2.0y = 3.0 \end{cases}$$

Avertissement numérique : le mauvais pivot

On considère toujours le même système linéaire où $\varepsilon \approx 0$.

$$\begin{cases} \varepsilon x + 1.0y = 1.0 \\ x + 2.0y = 3.0 \end{cases}$$

Supposons que l'on utilise (maladroitement) comme pivot ε :

$$\begin{array}{l} \text{On obtient} \\ \text{puis} \end{array} \begin{cases} x + \frac{1.0}{\varepsilon}y = \frac{1.0}{\varepsilon} \\ x + 2.0y = 3.0 \\ x + \frac{1.0}{\varepsilon}y = \frac{1.0}{\varepsilon} \\ (2.0 - \frac{1.0}{\varepsilon})y = (3.0 - \frac{1.0}{\varepsilon}) \end{cases}$$

Comment calculer l'inverse d'une matrice ?

L'algorithme de Gauss résout des systèmes d'équations linéaires.

Comment calculer l'inverse d'une matrice ?

L'algorithme de Gauss résout des systèmes d'équations linéaires.
Il calcule aussi efficacement l'inverse d'une matrice $A \in \mathbb{K}^{n \times n}$:

Comment calculer l'inverse d'une matrice ?

L'algorithme de Gauss résout des systèmes d'équations linéaires.
Il calcule aussi efficacement l'inverse d'une matrice $A \in \mathbb{K}^{n \times n}$:

- 1 On accole la matrice identité $1_{n \times n}$ à droite, noté $(A \mid 1_{n \times 1})$.

Comment calculer l'inverse d'une matrice ?

L'algorithme de Gauss résout des systèmes d'équations linéaires. Il calcule aussi efficacement l'inverse d'une matrice $A \in \mathbb{K}^{n \times n}$:

- 1 On accole la matrice identité $1_{n \times n}$ à droite, noté $(A \mid 1_{n \times 1})$.
- 2 On effectue la réduction de A sous forme échelonnée.
(Multiplication par des matrices inversibles élémentaires.)

Comment calculer l'inverse d'une matrice ?

L'algorithme de Gauss résout des systèmes d'équations linéaires.
Il calcule aussi efficacement l'inverse d'une matrice $A \in \mathbb{K}^{n \times n}$:

- 1 On accole la matrice identité $1_{n \times n}$ à droite, noté $(A \mid 1_{n \times 1})$.
- 2 On effectue la réduction de A sous forme échelonnée.
(Multiplication par des matrices inversibles élémentaires.)
- 3 Si A est inversible, alors on obtient à droite la matrice inverse.
(Sinon le calcul s'arrête et signale que A n'était pas inversible.)

Comment calculer l'inverse d'une matrice ?

L'algorithme de Gauss résout des systèmes d'équations linéaires. Il calcule aussi efficacement l'inverse d'une matrice $A \in \mathbb{K}^{n \times n}$:

- 1 On accole la matrice identité $1_{n \times n}$ à droite, noté $(A \mid 1_{n \times 1})$.
- 2 On effectue la réduction de A sous forme échelonnée.
(Multiplication par des matrices inversibles élémentaires.)
- 3 Si A est inversible, alors on obtient à droite la matrice inverse.
(Sinon le calcul s'arrête et signale que A n'était pas inversible.)

Pourquoi cet algorithme est-il correct ?

Comment calculer l'inverse d'une matrice ?

L'algorithme de Gauss résout des systèmes d'équations linéaires. Il calcule aussi efficacement l'inverse d'une matrice $A \in \mathbb{K}^{n \times n}$:

- 1 On accole la matrice identité $1_{n \times n}$ à droite, noté $(A \mid 1_{n \times 1})$.
- 2 On effectue la réduction de A sous forme échelonnée.
(Multiplication par des matrices inversibles élémentaires.)
- 3 Si A est inversible, alors on obtient à droite la matrice inverse.
(Sinon le calcul s'arrête et signale que A n'était pas inversible.)

Pourquoi cet algorithme est-il correct ?

- On commence par $(A_0 \mid B_0) = (A \mid 1_{n \times n})$.
Cette initialisation assure l'égalité $A_0 = B_0 A$.

Comment calculer l'inverse d'une matrice ?

L'algorithme de Gauss résout des systèmes d'équations linéaires. Il calcule aussi efficacement l'inverse d'une matrice $A \in \mathbb{K}^{n \times n}$:

- 1 On accole la matrice identité $1_{n \times n}$ à droite, noté $(A \mid 1_{n \times 1})$.
- 2 On effectue la réduction de A sous forme échelonnée.
(Multiplication par des matrices inversibles élémentaires.)
- 3 Si A est inversible, alors on obtient à droite la matrice inverse.
(Sinon le calcul s'arrête et signale que A n'était pas inversible.)

Pourquoi cet algorithme est-il correct ?

- On commence par $(A_0 \mid B_0) = (A \mid 1_{n \times n})$.
Cette initialisation assure l'égalité $A_0 = B_0 A$.
- Chaque opération élémentaire sur les lignes transforme $(A_k \mid B_k)$ en $(A_{k+1} \mid B_{k+1})$ où $A_{k+1} = T_k A_k$ et $B_{k+1} = T_k B_k$.

Comment calculer l'inverse d'une matrice ?

L'algorithme de Gauss résout des systèmes d'équations linéaires. Il calcule aussi efficacement l'inverse d'une matrice $A \in \mathbb{K}^{n \times n}$:

- 1 On accole la matrice identité $1_{n \times n}$ à droite, noté $(A \mid 1_{n \times 1})$.
- 2 On effectue la réduction de A sous forme échelonnée. (Multiplication par des matrices inversibles élémentaires.)
- 3 Si A est inversible, alors on obtient à droite la matrice inverse. (Sinon le calcul s'arrête et signale que A n'était pas inversible.)

Pourquoi cet algorithme est-il correct ?

- On commence par $(A_0 \mid B_0) = (A \mid 1_{n \times n})$. Cette initialisation assure l'égalité $A_0 = B_0 A$.
- Chaque opération élémentaire sur les lignes transforme $(A_k \mid B_k)$ en $(A_{k+1} \mid B_{k+1})$ où $A_{k+1} = T_k A_k$ et $B_{k+1} = T_k B_k$. On préserve ainsi l'égalité $A_k = B_k A$.

Comment calculer l'inverse d'une matrice ?

L'algorithme de Gauss résout des systèmes d'équations linéaires. Il calcule aussi efficacement l'inverse d'une matrice $A \in \mathbb{K}^{n \times n}$:

- 1 On accole la matrice identité $1_{n \times n}$ à droite, noté $(A \mid 1_{n \times 1})$.
- 2 On effectue la réduction de A sous forme échelonnée. (Multiplication par des matrices inversibles élémentaires.)
- 3 Si A est inversible, alors on obtient à droite la matrice inverse. (Sinon le calcul s'arrête et signale que A n'était pas inversible.)

Pourquoi cet algorithme est-il correct ?

- On commence par $(A_0 \mid B_0) = (A \mid 1_{n \times n})$. Cette initialisation assure l'égalité $A_0 = B_0 A$.
- Chaque opération élémentaire sur les lignes transforme $(A_k \mid B_k)$ en $(A_{k+1} \mid B_{k+1})$ où $A_{k+1} = T_k A_k$ et $B_{k+1} = T_k B_k$. On préserve ainsi l'égalité $A_k = B_k A$.
- À la fin on obtient $(A_m \mid B_m)$ où $A_m = \mathbf{1}$.

Comment calculer l'inverse d'une matrice ?


L'algorithme de Gauss résout des systèmes d'équations linéaires.
Il calcule aussi efficacement l'inverse d'une matrice $A \in \mathbb{K}^{n \times n}$:

- 1 On accole la matrice identité $1_{n \times n}$ à droite, noté $(A \mid 1_{n \times 1})$.
- 2 On effectue la réduction de A sous forme échelonnée.
(Multiplication par des matrices inversibles élémentaires.)
- 3 Si A est inversible, alors on obtient à droite la matrice inverse.
(Sinon le calcul s'arrête et signale que A n'était pas inversible.)

Pourquoi cet algorithme est-il correct ?

- On commence par $(A_0 \mid B_0) = (A \mid 1_{n \times n})$.
Cette initialisation assure l'égalité $A_0 = B_0 A$.
- Chaque opération élémentaire sur les lignes transforme $(A_k \mid B_k)$ en $(A_{k+1} \mid B_{k+1})$ où $A_{k+1} = T_k A_k$ et $B_{k+1} = T_k B_k$.
On préserve ainsi l'égalité $A_k = B_k A$.
- À la fin on obtient $(A_m \mid B_m)$ où $A_m = \mathbf{1}$.
Ainsi l'égalité $\mathbf{1} = B_m A$ assure que $B_m = A^{-1}$.


Comment calculer le déterminant d'une matrice ?

 Aussi élégante qu'elle soit, la belle formule

$$\det A = \sum_{\sigma} \text{sign}(\sigma) \cdot a_{1,\sigma(1)} \cdot a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}$$


n'est utilisable que pour n très petit. (Calculez 10! puis 20! puis 50!)

Comment calculer le déterminant d'une matrice ?


 Aussi élégante qu'elle soit, la belle formule

$$\det A = \sum_{\sigma} \text{sign}(\sigma) \cdot a_{1,\sigma(1)} \cdot a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}$$

n'est utilisable que pour n très petit. (Calculez 10! puis 20! puis 50!)


 De même, le développement par lignes et/ou colonnes n'est profitable que si la matrice présente beaucoup de zéros.

Comment calculer le déterminant d'une matrice ?


 Aussi élégante qu'elle soit, la belle formule

$$\det A = \sum_{\sigma} \text{sign}(\sigma) \cdot a_{1,\sigma(1)} \cdot a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}$$

n'est utilisable que pour n très petit. (Calculez 10! puis 20! puis 50!)


 De même, le développement par lignes et/ou colonnes n'est profitable que si la matrice présente beaucoup de zéros. Sinon on retombe sur la formule ci-dessus, trop lourde pour n grand.


Comment calculer le déterminant d'une matrice ?

 Aussi élégante qu'elle soit, la belle formule


$$\det A = \sum_{\sigma} \text{sign}(\sigma) \cdot a_{1,\sigma(1)} \cdot a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}$$

n'est utilisable que pour n très petit. (Calculez 10! puis 20! puis 50!)

 De même, le développement par lignes et/ou colonnes n'est profitable que si la matrice présente beaucoup de zéros. Sinon on retombe sur la formule ci-dessus, trop lourde pour n grand.


 À nouveau l'algorithme de Gauss vient à notre secours :
On réduit la matrice $A \in \mathbb{K}^{n \times n}$ à la matrice identité, où $\det \mathbf{1} = 1$.


Comment calculer le déterminant d'une matrice ?

 Aussi élégante qu'elle soit, la belle formule


$$\det A = \sum_{\sigma} \text{sign}(\sigma) \cdot a_{1,\sigma(1)} \cdot a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}$$

n'est utilisable que pour n très petit. (Calculez 10! puis 20! puis 50!)

 De même, le développement par lignes et/ou colonnes n'est profitable que si la matrice présente beaucoup de zéros. Sinon on retombe sur la formule ci-dessus, trop lourde pour n grand.


 À nouveau l'algorithme de Gauss vient à notre secours :
On réduit la matrice $A \in \mathbb{K}^{n \times n}$ à la matrice identité, où $\det \mathbf{1} = 1$.
Lors de l'algorithme on note le changement du déterminant :


Comment calculer le déterminant d'une matrice ?

 Aussi élégante qu'elle soit, la belle formule

$$\det A = \sum_{\sigma} \text{sign}(\sigma) \cdot a_{1,\sigma(1)} \cdot a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}$$


n'est utilisable que pour n très petit. (Calculez 10! puis 20! puis 50!)

 De même, le développement par lignes et/ou colonnes n'est profitable que si la matrice présente beaucoup de zéros. Sinon on retombe sur la formule ci-dessus, trop lourde pour n grand.

 À nouveau l'algorithme de Gauss vient à notre secours : On réduit la matrice $A \in \mathbb{K}^{n \times n}$ à la matrice identité, où $\det \mathbf{1} = 1$. Lors de l'algorithme on note le changement du déterminant :


- Chaque inversion $L_i \leftrightarrow L_j$ reverse le signe de $\det A$.


Comment calculer le déterminant d'une matrice ?

 Aussi élégante qu'elle soit, la belle formule

$$\det A = \sum_{\sigma} \text{sign}(\sigma) \cdot a_{1,\sigma(1)} \cdot a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}$$


n'est utilisable que pour n très petit. (Calculez 10! puis 20! puis 50!)

 De même, le développement par lignes et/ou colonnes n'est profitable que si la matrice présente beaucoup de zéros. Sinon on retombe sur la formule ci-dessus, trop lourde pour n grand.

 À nouveau l'algorithme de Gauss vient à notre secours : On réduit la matrice $A \in \mathbb{K}^{n \times n}$ à la matrice identité, où $\det \mathbf{1} = 1$. Lors de l'algorithme on note le changement du déterminant :


- Chaque inversion $L_i \leftrightarrow L_j$ renverse le signe de $\det A$.
- Chaque multiplication $L_i \leftarrow \lambda L_i$ multiplie $\det A$ par λ .


Comment calculer le déterminant d'une matrice ?

 Aussi élégante qu'elle soit, la belle formule

$$\det A = \sum_{\sigma} \text{sign}(\sigma) \cdot a_{1,\sigma(1)} \cdot a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}$$

n'est utilisable que pour n très petit. (Calculez 10! puis 20! puis 50!)

 De même, le développement par lignes et/ou colonnes n'est profitable que si la matrice présente beaucoup de zéros. Sinon on retombe sur la formule ci-dessus, trop lourde pour n grand.

 À nouveau l'algorithme de Gauss vient à notre secours : On réduit la matrice $A \in \mathbb{K}^{n \times n}$ à la matrice identité, où $\det \mathbf{1} = 1$. Lors de l'algorithme on note le changement du déterminant :

- Chaque inversion $L_i \leftrightarrow L_j$ renverse le signe de $\det A$.
- Chaque multiplication $L_i \leftarrow \lambda L_i$ multiplie $\det A$ par λ .
- Chaque addition $L_i \leftarrow L_i + \lambda L_j$ laisse $\det A$ invariant.

Avertissement numérique

Exemple. On considère l'équation $Ax = y$ avec

$$A = \begin{pmatrix} 0.780 & 0.563 \\ 0.913 & 0.659 \end{pmatrix} \quad \text{et} \quad y = \begin{pmatrix} 0.217 \\ 0.254 \end{pmatrix}.$$

Avertissement numérique

Exemple. On considère l'équation $Ax = y$ avec

$$A = \begin{pmatrix} 0.780 & 0.563 \\ 0.913 & 0.659 \end{pmatrix} \quad \text{et} \quad y = \begin{pmatrix} 0.217 \\ 0.254 \end{pmatrix}.$$

La matrice A est inversible, car on trouve $\det A = 10^{-6}$.

Avertissement numérique

Exemple. On considère l'équation $Ax = y$ avec

$$A = \begin{pmatrix} 0.780 & 0.563 \\ 0.913 & 0.659 \end{pmatrix} \quad \text{et} \quad y = \begin{pmatrix} 0.217 \\ 0.254 \end{pmatrix}.$$

La matrice A est inversible, car on trouve $\det A = 10^{-6}$.

Lequel des deux résultats approchés suivants est meilleur :

$$\tilde{x} = \begin{pmatrix} 0.999 \\ -1.001 \end{pmatrix} \quad \text{ou} \quad \hat{x} = \begin{pmatrix} 0.341 \\ -0.087 \end{pmatrix}?$$

Avertissement numérique

Exemple. On considère l'équation $Ax = y$ avec

$$A = \begin{pmatrix} 0.780 & 0.563 \\ 0.913 & 0.659 \end{pmatrix} \quad \text{et} \quad y = \begin{pmatrix} 0.217 \\ 0.254 \end{pmatrix}.$$

La matrice A est inversible, car on trouve $\det A = 10^{-6}$.

Lequel des deux résultats approchés suivants est meilleur :

$$\tilde{x} = \begin{pmatrix} 0.999 \\ -1.001 \end{pmatrix} \quad \text{ou} \quad \hat{x} = \begin{pmatrix} 0.341 \\ -0.087 \end{pmatrix}?$$

On pourrait naïvement calculer les erreurs $|A\tilde{x} - y|$ et $|A\hat{x} - y|$, puis choisir la solution qui minimise cette erreur.

Avertissement numérique

Exemple. On considère l'équation $Ax = y$ avec

$$A = \begin{pmatrix} 0.780 & 0.563 \\ 0.913 & 0.659 \end{pmatrix} \quad \text{et} \quad y = \begin{pmatrix} 0.217 \\ 0.254 \end{pmatrix}.$$

La matrice A est inversible, car on trouve $\det A = 10^{-6}$.

Lequel des deux résultats approchés suivants est meilleur :

$$\tilde{x} = \begin{pmatrix} 0.999 \\ -1.001 \end{pmatrix} \quad \text{ou} \quad \hat{x} = \begin{pmatrix} 0.341 \\ -0.087 \end{pmatrix}?$$

On pourrait naïvement calculer les erreurs $|A\tilde{x} - y|$ et $|A\hat{x} - y|$, puis choisir la solution qui minimise cette erreur. Ici c'est \hat{x} , car

$$A\tilde{x} - y = \begin{pmatrix} -0.001343 \\ -0.001572 \end{pmatrix} \quad \text{et} \quad A\hat{x} - y = \begin{pmatrix} -0.0000001 \\ 0.0000000 \end{pmatrix}.$$

Avertissement numérique

Exemple. On considère l'équation $Ax = y$ avec

$$A = \begin{pmatrix} 0.780 & 0.563 \\ 0.913 & 0.659 \end{pmatrix} \quad \text{et} \quad y = \begin{pmatrix} 0.217 \\ 0.254 \end{pmatrix}.$$

La matrice A est inversible, car on trouve $\det A = 10^{-6}$.

Lequel des deux résultats approchés suivants est meilleur :

$$\tilde{x} = \begin{pmatrix} 0.999 \\ -1.001 \end{pmatrix} \quad \text{ou} \quad \hat{x} = \begin{pmatrix} 0.341 \\ -0.087 \end{pmatrix}?$$

On pourrait naïvement calculer les erreurs $|A\tilde{x} - y|$ et $|A\hat{x} - y|$, puis choisir la solution qui minimise cette erreur. Ici c'est \hat{x} , car

$$A\tilde{x} - y = \begin{pmatrix} -0.001343 \\ -0.001572 \end{pmatrix} \quad \text{et} \quad A\hat{x} - y = \begin{pmatrix} -0.0000001 \\ 0.0000000 \end{pmatrix}.$$

Pourtant la solution exacte est $x = \begin{pmatrix} +1 \\ -1 \end{pmatrix}$... Le vérifier !

Avertissement numérique

Exemple. On considère l'équation $Ax = y$ avec

$$A = \begin{pmatrix} 0.780 & 0.563 \\ 0.913 & 0.659 \end{pmatrix} \quad \text{et} \quad y = \begin{pmatrix} 0.217 \\ 0.254 \end{pmatrix}.$$

La matrice A est inversible, car on trouve $\det A = 10^{-6}$.

Lequel des deux résultats approchés suivants est meilleur :

$$\tilde{x} = \begin{pmatrix} 0.999 \\ -1.001 \end{pmatrix} \quad \text{ou} \quad \hat{x} = \begin{pmatrix} 0.341 \\ -0.087 \end{pmatrix}?$$

On pourrait naïvement calculer les erreurs $|A\tilde{x} - y|$ et $|A\hat{x} - y|$, puis choisir la solution qui minimise cette erreur. Ici c'est \hat{x} , car

$$A\tilde{x} - y = \begin{pmatrix} -0.001343 \\ -0.001572 \end{pmatrix} \quad \text{et} \quad A\hat{x} - y = \begin{pmatrix} -0.0000001 \\ 0.0000000 \end{pmatrix}.$$

Pourtant la solution exacte est $x = \begin{pmatrix} +1 \\ -1 \end{pmatrix}$... Le vérifier !

À notre grande surprise, c'est donc \tilde{x} qui est le plus proche.

Avertissement numérique

Exemple. On considère l'équation $Ax = y$ avec

$$A = \begin{pmatrix} 0.780 & 0.563 \\ 0.913 & 0.659 \end{pmatrix} \quad \text{et} \quad y = \begin{pmatrix} 0.217 \\ 0.254 \end{pmatrix}.$$

La matrice A est inversible, car on trouve $\det A = 10^{-6}$.

Lequel des deux résultats approchés suivants est meilleur :

$$\tilde{x} = \begin{pmatrix} 0.999 \\ -1.001 \end{pmatrix} \quad \text{ou} \quad \hat{x} = \begin{pmatrix} 0.341 \\ -0.087 \end{pmatrix}?$$

On pourrait naïvement calculer les erreurs $|A\tilde{x} - y|$ et $|A\hat{x} - y|$, puis choisir la solution qui minimise cette erreur. Ici c'est \hat{x} , car

$$A\tilde{x} - y = \begin{pmatrix} -0.001343 \\ -0.001572 \end{pmatrix} \quad \text{et} \quad A\hat{x} - y = \begin{pmatrix} -0.0000001 \\ 0.0000000 \end{pmatrix}.$$

Pourtant la solution exacte est $x = \begin{pmatrix} +1 \\ -1 \end{pmatrix}$... Le vérifier !

À notre grande surprise, c'est donc \tilde{x} qui est le plus proche.
Comment expliquer puis quantifier ce phénomène étrange ?

La norme d'une matrice

Définition

Pour $A \in \mathbb{C}^{n \times n}$ on définit la *norme* par $\|A\| := \sup\{ |Ax|; |x| \leq 1 \}$.

La norme d'une matrice

Définition

Pour $A \in \mathbb{C}^{n \times n}$ on définit la *norme* par $\|A\| := \sup\{|Ax|; |x| \leq 1\}$.

La norme $\|A\|$ mesure la « distorsion » : on a $|Ax| \leq \|A\| \cdot |x|$ pour tout $x \in \mathbb{C}^n$, et $\|A\|$ est la plus petite valeur assurant cette inégalité.

La norme d'une matrice

Définition

Pour $A \in \mathbb{C}^{n \times n}$ on définit la *norme* par $\|A\| := \sup\{|Ax|; |x| \leq 1\}$.

La norme $\|A\|$ mesure la « distorsion » : on a $|Ax| \leq \|A\| \cdot |x|$ pour tout $x \in \mathbb{C}^n$, et $\|A\|$ est la plus petite valeur assurant cette inégalité.

Exemple : Si $A \sim \text{diag}(\lambda_1, \dots, \lambda_n)$, alors $\|A\| = \max\{|\lambda_1|, \dots, |\lambda_n|\}$.

La norme d'une matrice

Définition

Pour $A \in \mathbb{C}^{n \times n}$ on définit la *norme* par $\|A\| := \sup\{|Ax|; |x| \leq 1\}$.

La norme $\|A\|$ mesure la « distorsion » : on a $|Ax| \leq \|A\| \cdot |x|$ pour tout $x \in \mathbb{C}^n$, et $\|A\|$ est la plus petite valeur assurant cette inégalité.

Exemple : Si $A \sim \text{diag}(\lambda_1, \dots, \lambda_n)$, alors $\|A\| = \max\{|\lambda_1|, \dots, |\lambda_n|\}$.

Proposition

L'application $A \mapsto \|A\|$ définit une norme sur $\mathbb{C}^{n \times n}$:

homogénéité : $\|\lambda A\| = |\lambda| \cdot \|A\|$,

inégalité triangulaire : $\|A + B\| \leq \|A\| + \|B\|$,

positivité : $\|A\| \geq 0$, et $\|A\| = 0 \Leftrightarrow A = 0$.

Le conditionnement d'une matrice

Définition

On appelle $\text{cond}(A) := \|A\| \cdot \|A^{-1}\|$ le *conditionnement* de A .

Le conditionnement mesure la sensibilité aux erreurs

On considère une matrice A inversible et une solution $Ax = y$.

Le conditionnement mesure la sensibilité aux erreurs

On considère une matrice A inversible et une solution $Ax = y$.

Si $A\tilde{x} = \tilde{y}$, où $\tilde{y} = y + \delta y$ entraîne $\tilde{x} = x + \delta x$, alors

$$\frac{1}{\text{cond}(A)} \frac{|\delta y|}{|y|} \leq \frac{|\delta x|}{|x|} \leq \text{cond}(A) \frac{|\delta y|}{|y|}.$$

Conditionnement d'une matrice : exemple

Considérons la matrice $A = \begin{pmatrix} 7 & 10 \\ 5 & 7 \end{pmatrix}$ qui a pour inverse $A^{-1} = \begin{pmatrix} -7 & 10 \\ 5 & -7 \end{pmatrix}$.

Conditionnement d'une matrice : exemple

Considérons la matrice $A = \begin{pmatrix} 7 & 10 \\ 5 & 7 \end{pmatrix}$ qui a pour inverse $A^{-1} = \begin{pmatrix} -7 & 10 \\ 5 & -7 \end{pmatrix}$.

Pour les valeurs propres de A on trouve $\lambda_1 = 7 - 5\sqrt{2} \approx -0.071$
et $\lambda_2 = 7 + 5\sqrt{2} \approx 14.071$, donc la norme de A est $\|A\| \approx 14.071$.

Conditionnement d'une matrice : exemple

Considérons la matrice $A = \begin{pmatrix} 7 & 10 \\ 5 & 7 \end{pmatrix}$ qui a pour inverse $A^{-1} = \begin{pmatrix} -7 & 10 \\ 5 & -7 \end{pmatrix}$.

Pour les valeurs propres de A on trouve $\lambda_1 = 7 - 5\sqrt{2} \approx -0.071$
et $\lambda_2 = 7 + 5\sqrt{2} \approx 14.071$, donc la norme de A est $\|A\| \approx 14.071$.

Les valeurs propres de A^{-1} sont $-7 \pm 5\sqrt{2}$, donc $\|A^{-1}\| \approx 14.071$.

Ainsi le conditionnement de A est $\text{cond}(A) = \|A\| \cdot \|A^{-1}\| \approx 198$.

Ceci indique que $Ax = y$ peut être sensible aux perturbations de y .

Conditionnement d'une matrice : exemple

Considérons la matrice $A = \begin{pmatrix} 7 & 10 \\ 5 & 7 \end{pmatrix}$ qui a pour inverse $A^{-1} = \begin{pmatrix} -7 & 10 \\ 5 & -7 \end{pmatrix}$.

Pour les valeurs propres de A on trouve $\lambda_1 = 7 - 5\sqrt{2} \approx -0.071$
et $\lambda_2 = 7 + 5\sqrt{2} \approx 14.071$, donc la norme de A est $\|A\| \approx 14.071$.

Les valeurs propres de A^{-1} sont $-7 \pm 5\sqrt{2}$, donc $\|A^{-1}\| \approx 14.071$.

Ainsi le conditionnement de A est $\text{cond}(A) = \|A\| \cdot \|A^{-1}\| \approx 198$.

Ceci indique que $Ax = y$ peut être sensible aux perturbations de y .

Exemple. $Ax = \begin{pmatrix} 1.00 \\ 0.70 \end{pmatrix}$ donne $x = \begin{pmatrix} 0.00 \\ 0.10 \end{pmatrix}$,

alors que $A\tilde{x} = \begin{pmatrix} 1.01 \\ 0.69 \end{pmatrix}$ donne $\tilde{x} = \begin{pmatrix} -0.17 \\ 0.22 \end{pmatrix}$.

Conditionnement d'une matrice : exemple

Considérons la matrice $A = \begin{pmatrix} 7 & 10 \\ 5 & 7 \end{pmatrix}$ qui a pour inverse $A^{-1} = \begin{pmatrix} -7 & 10 \\ 5 & -7 \end{pmatrix}$.

Pour les valeurs propres de A on trouve $\lambda_1 = 7 - 5\sqrt{2} \approx -0.071$ et $\lambda_2 = 7 + 5\sqrt{2} \approx 14.071$, donc la norme de A est $\|A\| \approx 14.071$.

Les valeurs propres de A^{-1} sont $-7 \pm 5\sqrt{2}$, donc $\|A^{-1}\| \approx 14.071$.

Ainsi le conditionnement de A est $\text{cond}(A) = \|A\| \cdot \|A^{-1}\| \approx 198$.

Ceci indique que $Ax = y$ peut être sensible aux perturbations de y .

Exemple. $Ax = \begin{pmatrix} 1.00 \\ 0.70 \end{pmatrix}$ donne $x = \begin{pmatrix} 0.00 \\ 0.10 \end{pmatrix}$,

alors que $A\tilde{x} = \begin{pmatrix} 1.01 \\ 0.69 \end{pmatrix}$ donne $\tilde{x} = \begin{pmatrix} -0.17 \\ 0.22 \end{pmatrix}$.

Conclusion

Un conditionnement $\text{cond}(A) \approx 10^c$ indique une possible perte de précision, typiquement de c décimales : En général, la donnée de y avec une précision de ℓ décimales significatives mène à une solution x avec une précision de $\ell - c$ décimales significatives seulement.

Sommaire

- 1 Résolution de systèmes d'équations linéaires
- 2 Réduction des endomorphismes
 - Espaces vectoriels et applications linéaires
 - Vecteurs propres, polynôme caractéristique
 - Polynôme minimal, méthodes de calcul
- 3 Méthodes approchées itératives
- 4 Comment fonctionne Google ?

Définition

Soit $A \in \mathbb{K}^{n \times n}$ une matrice carrée.

Vecteurs et valeurs propres, polynôme caractéristique

Définition

Soit $A \in \mathbb{K}^{n \times n}$ une matrice carrée. Un vecteur non nul $v \in \mathbb{K}^n \setminus \{0\}$ est un *vecteur propre* de A si $Av = \lambda v$ où $\lambda \in \mathbb{K}$.

Vecteurs et valeurs propres, polynôme caractéristique

Définition

Soit $A \in \mathbb{K}^{n \times n}$ une matrice carrée. Un vecteur non nul $v \in \mathbb{K}^n \setminus \{0\}$ est un *vecteur propre* de A si $Av = \lambda v$ où $\lambda \in \mathbb{K}$. Dans ce cas λ est une *valeur propre* de A , et que v est un vecteur propre associé à λ .

Vecteurs et valeurs propres, polynôme caractéristique

Définition

Soit $A \in \mathbb{K}^{n \times n}$ une matrice carrée. Un vecteur non nul $v \in \mathbb{K}^n \setminus \{0\}$ est un *vecteur propre* de A si $Av = \lambda v$ où $\lambda \in \mathbb{K}$. Dans ce cas λ est une *valeur propre* de A , et que v est un vecteur propre associé à λ .

Puisque $Av = \lambda v$ équivaut à $(A - \lambda I)v = 0$, on cherche des valeurs λ pour lesquelles la matrice $A - \lambda I$ ait un noyau non trivial.

Vecteurs et valeurs propres, polynôme caractéristique

Définition

Soit $A \in \mathbb{K}^{n \times n}$ une matrice carrée. Un vecteur non nul $v \in \mathbb{K}^n \setminus \{0\}$ est un *vecteur propre* de A si $Av = \lambda v$ où $\lambda \in \mathbb{K}$. Dans ce cas λ est une *valeur propre* de A , et que v est un vecteur propre associé à λ .

Puisque $Av = \lambda v$ équivaut à $(A - \lambda I)v = 0$, on cherche des valeurs λ pour lesquelles la matrice $A - \lambda I$ ait un noyau non trivial.

Définition

On appelle $\chi_A := \det(A - XI)$ le *polynôme caractéristique* de A .

Vecteurs et valeurs propres, polynôme caractéristique

Définition

Soit $A \in \mathbb{K}^{n \times n}$ une matrice carrée. Un vecteur non nul $v \in \mathbb{K}^n \setminus \{0\}$ est un *vecteur propre* de A si $Av = \lambda v$ où $\lambda \in \mathbb{K}$. Dans ce cas λ est une *valeur propre* de A , et que v est un vecteur propre associé à λ .

Puisque $Av = \lambda v$ équivaut à $(A - \lambda I)v = 0$, on cherche des valeurs λ pour lesquelles la matrice $A - \lambda I$ ait un noyau non trivial.

Définition

On appelle $\chi_A := \det(A - XI)$ le *polynôme caractéristique* de A .

C'est effectivement un polynôme dans $\mathbb{K}[X]$ de degré n .

Vecteurs et valeurs propres, polynôme caractéristique

Définition

Soit $A \in \mathbb{K}^{n \times n}$ une matrice carrée. Un vecteur non nul $v \in \mathbb{K}^n \setminus \{0\}$ est un *vecteur propre* de A si $Av = \lambda v$ où $\lambda \in \mathbb{K}$. Dans ce cas λ est une *valeur propre* de A , et que v est un vecteur propre associé à λ .

Puisque $Av = \lambda v$ équivaut à $(A - \lambda I)v = 0$, on cherche des valeurs λ pour lesquelles la matrice $A - \lambda I$ ait un noyau non trivial.

Définition

On appelle $\chi_A := \det(A - XI)$ le *polynôme caractéristique* de A .

C'est effectivement un polynôme dans $\mathbb{K}[X]$ de degré n . On a $\chi_A(\lambda) = 0$ si et seulement si λ est une valeur propre de A .

Vecteurs et valeurs propres, polynôme caractéristique

Définition

Soit $A \in \mathbb{K}^{n \times n}$ une matrice carrée. Un vecteur non nul $v \in \mathbb{K}^n \setminus \{0\}$ est un *vecteur propre* de A si $Av = \lambda v$ où $\lambda \in \mathbb{K}$. Dans ce cas λ est une *valeur propre* de A , et que v est un vecteur propre associé à λ .

Puisque $Av = \lambda v$ équivaut à $(A - \lambda I)v = 0$, on cherche des valeurs λ pour lesquelles la matrice $A - \lambda I$ ait un noyau non trivial.

Définition

On appelle $\chi_A := \det(A - XI)$ le *polynôme caractéristique* de A .

C'est effectivement un polynôme dans $\mathbb{K}[X]$ de degré n . On a $\chi_A(\lambda) = 0$ si et seulement si λ est une valeur propre de A .

◆ Pour trouver les valeurs propres de A on est ainsi ramené à résoudre une équation polynomiale, comme déjà discuté.

Vecteurs et valeurs propres, polynôme caractéristique

Définition

Soit $A \in \mathbb{K}^{n \times n}$ une matrice carrée. Un vecteur non nul $v \in \mathbb{K}^n \setminus \{0\}$ est un *vecteur propre* de A si $Av = \lambda v$ où $\lambda \in \mathbb{K}$. Dans ce cas λ est une *valeur propre* de A , et que v est un vecteur propre associé à λ .

Puisque $Av = \lambda v$ équivaut à $(A - \lambda I)v = 0$, on cherche des valeurs λ pour lesquelles la matrice $A - \lambda I$ ait un noyau non trivial.

Définition

On appelle $\chi_A := \det(A - XI)$ le *polynôme caractéristique* de A .

C'est effectivement un polynôme dans $\mathbb{K}[X]$ de degré n . On a $\chi_A(\lambda) = 0$ si et seulement si λ est une valeur propre de A .

◆ Pour trouver les valeurs propres de A on est ainsi ramené à résoudre une équation polynomiale, comme déjà discuté.
Vive la résolution d'équations polynomiales !

Vecteurs et valeurs propres, polynôme caractéristique

Définition

Soit $A \in \mathbb{K}^{n \times n}$ une matrice carrée. Un vecteur non nul $v \in \mathbb{K}^n \setminus \{0\}$ est un *vecteur propre* de A si $Av = \lambda v$ où $\lambda \in \mathbb{K}$. Dans ce cas λ est une *valeur propre* de A , et que v est un vecteur propre associé à λ .

Puisque $Av = \lambda v$ équivaut à $(A - \lambda I)v = 0$, on cherche des valeurs λ pour lesquelles la matrice $A - \lambda I$ ait un noyau non trivial.

Définition

On appelle $\chi_A := \det(A - XI)$ le *polynôme caractéristique* de A .

C'est effectivement un polynôme dans $\mathbb{K}[X]$ de degré n . On a $\chi_A(\lambda) = 0$ si et seulement si λ est une valeur propre de A .

◆ Pour trouver les valeurs propres de A on est ainsi ramené à résoudre une équation polynomiale, comme déjà discuté.
Vive la résolution d'équations polynomiales !

◆ Pour trouver l'espace des vecteurs propres associés à λ on est ramené à trouver le noyau de la matrice $A - \lambda I$, comme déjà discuté.

Vecteurs et valeurs propres, polynôme caractéristique

Définition

Soit $A \in \mathbb{K}^{n \times n}$ une matrice carrée. Un vecteur non nul $v \in \mathbb{K}^n \setminus \{0\}$ est un *vecteur propre* de A si $Av = \lambda v$ où $\lambda \in \mathbb{K}$. Dans ce cas λ est une *valeur propre* de A , et que v est un vecteur propre associé à λ .

Puisque $Av = \lambda v$ équivaut à $(A - \lambda I)v = 0$, on cherche des valeurs λ pour lesquelles la matrice $A - \lambda I$ ait un noyau non trivial.

Définition

On appelle $\chi_A := \det(A - XI)$ le *polynôme caractéristique* de A .

C'est effectivement un polynôme dans $\mathbb{K}[X]$ de degré n . On a $\chi_A(\lambda) = 0$ si et seulement si λ est une valeur propre de A .

◆ Pour trouver les valeurs propres de A on est ainsi ramené à résoudre une équation polynomiale, comme déjà discuté.
Vive la résolution d'équations polynomiales !

◆ Pour trouver l'espace des vecteurs propres associés à λ on est ramené à trouver le noyau de la matrice $A - \lambda I$, comme déjà discuté.
Vive l'algorithme de Gauss !

Le polynôme minimal d'une matrice carrée

Soit $A \in \mathbb{K}^{n \times n}$ une matrice carrée sur un corps \mathbb{K} .

Le polynôme minimal d'une matrice carrée

Soit $A \in \mathbb{K}^{n \times n}$ une matrice carrée sur un corps \mathbb{K} .

Pour tout polynôme $P = a_0 + a_1X + \cdots + a_kX^k$ dans $\mathbb{K}[X]$
on définit la matrice $P(A) := a_0I + a_1A + \cdots + a_kA^k$ dans $\mathbb{K}^{n \times n}$.

Le polynôme minimal d'une matrice carrée

Soit $A \in \mathbb{K}^{n \times n}$ une matrice carrée sur un corps \mathbb{K} .

Pour tout polynôme $P = a_0 + a_1X + \cdots + a_kX^k$ dans $\mathbb{K}[X]$
on définit la matrice $P(A) := a_0I + a_1A + \cdots + a_kA^k$ dans $\mathbb{K}^{n \times n}$.

Définition

Le *polynôme minimal* d'une matrice carrée $A \in \mathbb{K}^{n \times n}$ est l'unique polynôme unitaire $\mu \in \mathbb{K}[X]$ de degré minimal tel que $\mu(A) = 0$.

Polynôme minimal : un premier exemple

Nous cherchons le polynôme minimal de la matrice $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$.

Polynôme minimal : un premier exemple

Nous cherchons le polynôme minimal de la matrice $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$.

D'abord on pourra contempler quelques puissances de A :

$$I = A^0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad A^1 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad A^2 = \begin{pmatrix} 7 & 10 \\ 15 & 22 \end{pmatrix}, \quad \dots$$

Polynôme minimal : un premier exemple

Nous cherchons le polynôme minimal de la matrice $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$.

D'abord on pourra contempler quelques puissances de A :

$$I = A^0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad A^1 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad A^2 = \begin{pmatrix} 7 & 10 \\ 15 & 22 \end{pmatrix}, \quad \dots$$

Degré 1 : existe-t-il un polynôme annulateur $P = a_0 + X$?

Polynôme minimal : un premier exemple

Nous cherchons le polynôme minimal de la matrice $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$.

D'abord on pourra contempler quelques puissances de A :

$$I = A^0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad A^1 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad A^2 = \begin{pmatrix} 7 & 10 \\ 15 & 22 \end{pmatrix}, \quad \dots$$

Degré 1 : existe-t-il un polynôme annulateur $P = a_0 + X$?

Ceci voudrait dire que $P(A) = a_0A^0 + A^1 = 0$. C'est impossible !

Polynôme minimal : un premier exemple

Nous cherchons le polynôme minimal de la matrice $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$.

D'abord on pourra contempler quelques puissances de A :

$$I = A^0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad A^1 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad A^2 = \begin{pmatrix} 7 & 10 \\ 15 & 22 \end{pmatrix}, \quad \dots$$

Degré 1 : existe-t-il un polynôme annulateur $P = a_0 + X$?

Ceci voudrait dire que $P(A) = a_0A^0 + A^1 = 0$. C'est impossible !

Degré 2 : existe-t-il un polynôme annulateur $P = a_0 + a_1X + X^2$?

Polynôme minimal : un premier exemple

Nous cherchons le polynôme minimal de la matrice $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$.

D'abord on pourra contempler quelques puissances de A :

$$I = A^0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad A^1 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad A^2 = \begin{pmatrix} 7 & 10 \\ 15 & 22 \end{pmatrix}, \quad \dots$$

Degré 1 : existe-t-il un polynôme annulateur $P = a_0 + X$?

Ceci voudrait dire que $P(A) = a_0A^0 + A^1 = 0$. C'est impossible !

Degré 2 : existe-t-il un polynôme annulateur $P = a_0 + a_1X + X^2$?

Ceci voudrait dire que $P(A) = a_0A^0 + a_1A^1 + A^2 = 0$. Calculons !

$$P(A) = \begin{pmatrix} a_0 + a_1 + 7 & 2a_1 + 10 \\ 3a_1 + 15 & a_0 + 4a_1 + 22 \end{pmatrix}$$

On voit que le polynôme $P = X^2 - 5X - 2$ annule la matrice A .

Polynôme minimal : un premier exemple

Nous cherchons le polynôme minimal de la matrice $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$.

D'abord on pourra contempler quelques puissances de A :

$$I = A^0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad A^1 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad A^2 = \begin{pmatrix} 7 & 10 \\ 15 & 22 \end{pmatrix}, \quad \dots$$

Degré 1 : existe-t-il un polynôme annulateur $P = a_0 + X$?

Ceci voudrait dire que $P(A) = a_0A^0 + A^1 = 0$. C'est impossible !

Degré 2 : existe-t-il un polynôme annulateur $P = a_0 + a_1X + X^2$?

Ceci voudrait dire que $P(A) = a_0A^0 + a_1A^1 + A^2 = 0$. Calculons !

$$P(A) = \begin{pmatrix} a_0 + a_1 + 7 & 2a_1 + 10 \\ 3a_1 + 15 & a_0 + 4a_1 + 22 \end{pmatrix}$$

On voit que le polynôme $P = X^2 - 5X - 2$ annule la matrice A .

C'est donc le polynôme minimal de A , car degré < 2 est impossible.

Polynôme minimal : un premier exemple

Nous cherchons le polynôme minimal de la matrice $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$.

D'abord on pourra contempler quelques puissances de A :

$$I = A^0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad A^1 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad A^2 = \begin{pmatrix} 7 & 10 \\ 15 & 22 \end{pmatrix}, \quad \dots$$

Degré 1 : existe-t-il un polynôme annulateur $P = a_0 + X$?

Ceci voudrait dire que $P(A) = a_0A^0 + A^1 = 0$. C'est impossible !

Degré 2 : existe-t-il un polynôme annulateur $P = a_0 + a_1X + X^2$?

Ceci voudrait dire que $P(A) = a_0A^0 + a_1A^1 + A^2 = 0$. Calculons !

$$P(A) = \begin{pmatrix} a_0 + a_1 + 7 & 2a_1 + 10 \\ 3a_1 + 15 & a_0 + 4a_1 + 22 \end{pmatrix}$$

On voit que le polynôme $P = X^2 - 5X - 2$ annule la matrice A .

C'est donc le polynôme minimal de A , car degré < 2 est impossible.

Remarquons finalement qu'ici le polynôme minimal μ_A coïncide avec le polynôme caractéristique $\chi_A = \det(A - XI) = X^2 - 5X - 2$.

Polynômes minimal : le théorème de Cayley-Hamilton

Théorème (de Cayley-Hamilton, admis)

Le polynôme caractéristique χ_A annule la matrice A .

Polynômes minimal : le théorème de Cayley-Hamilton

Théorème (de Cayley-Hamilton, admis)

*Le polynôme caractéristique χ_A annule la matrice A .
Autrement dit, le polynôme minimal μ_A divise χ_A .*



Polynômes minimal : le théorème de Cayley-Hamilton

Théorème (de Cayley-Hamilton, admis)

*Le polynôme caractéristique χ_A annule la matrice A .
Autrement dit, le polynôme minimal μ_A divise χ_A .*



En particulier on en déduit que $\deg \mu_A \leq \deg \chi_A = n$.

Polynômes minimal : le théorème de Cayley-Hamilton

Théorème (de Cayley-Hamilton, admis)

*Le polynôme caractéristique χ_A annule la matrice A .
Autrement dit, le polynôme minimal μ_A divise χ_A .* □

En particulier on en déduit que $\deg \mu_A \leq \deg \chi_A = n$.

Exemple : Pour $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ on trouve le polynôme caractéristique

$$\chi_A = -X^3 + 15X^2 + 18X.$$

Polynômes minimal : le théorème de Cayley-Hamilton

Théorème (de Cayley-Hamilton, admis)

Le polynôme caractéristique χ_A annule la matrice A .
Autrement dit, le polynôme minimal μ_A divise χ_A . □

En particulier on en déduit que $\deg \mu_A \leq \deg \chi_A = n$.

Exemple : Pour $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ on trouve le polynôme caractéristique

$$\chi_A = -X^3 + 15X^2 + 18X.$$

Les puissances de A sont

$$A^1 = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}, \quad A^2 = \begin{pmatrix} 30 & 36 & 42 \\ 66 & 81 & 96 \\ 102 & 126 & 150 \end{pmatrix}, \quad A^3 = \begin{pmatrix} 468 & 576 & 684 \\ 1062 & 1305 & 1548 \\ 1656 & 2034 & 2412 \end{pmatrix}.$$

Polynômes minimal : le théorème de Cayley-Hamilton

Théorème (de Cayley-Hamilton, admis)

Le polynôme caractéristique χ_A annule la matrice A .
Autrement dit, le polynôme minimal μ_A divise χ_A . □

En particulier on en déduit que $\deg \mu_A \leq \deg \chi_A = n$.

Exemple : Pour $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ on trouve le polynôme caractéristique

$$\chi_A = -X^3 + 15X^2 + 18X.$$

Les puissances de A sont

$$A^1 = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}, \quad A^2 = \begin{pmatrix} 30 & 36 & 42 \\ 66 & 81 & 96 \\ 102 & 126 & 150 \end{pmatrix}, \quad A^3 = \begin{pmatrix} 468 & 576 & 684 \\ 1062 & 1305 & 1548 \\ 1656 & 2034 & 2412 \end{pmatrix}.$$

On constate (après calcul) qu'effectivement

$$A^3 - 15A^2 - 18A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Polynômes minimal et caractéristique : exemples

Exemple où $A \in \mathbb{C}^{3 \times 3}$ a trois valeurs propres distinctes :

$$A = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} \quad \Rightarrow \quad \chi_A = \mu_A = (X - a)(X - b)(X - c)$$

Polynômes minimal et caractéristique : exemples

Exemple où $A \in \mathbb{C}^{3 \times 3}$ a trois valeurs propres distinctes :

$$A = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} \quad \Rightarrow \quad \chi_A = \mu_A = (X - a)(X - b)(X - c)$$

Exemples où $A \in \mathbb{C}^{3 \times 3}$ n'a que deux valeurs propres distinctes :

$$A = \begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{pmatrix} \quad \Rightarrow \quad \chi_A = (X - a)^2(X - b)$$

$$\mu_A = (X - a)(X - b)$$

$$A = \begin{pmatrix} a & 1 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{pmatrix} \quad \Rightarrow \quad \chi_A = \mu_A = (X - a)^2(X - b)$$

Polynômes minimal et caractéristique : exemples

Exemple où $A \in \mathbb{C}^{3 \times 3}$ a trois valeurs propres distinctes :

$$A = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} \Rightarrow \chi_A = \mu_A = (X - a)(X - b)(X - c)$$

Exemples où $A \in \mathbb{C}^{3 \times 3}$ n'a que deux valeurs propres distinctes :

$$A = \begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{pmatrix} \Rightarrow \begin{aligned} \chi_A &= (X - a)^2(X - b) \\ \mu_A &= (X - a)(X - b) \end{aligned}$$

$$A = \begin{pmatrix} a & 1 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{pmatrix} \Rightarrow \chi_A = \mu_A = (X - a)^2(X - b)$$

Exemples où $A \in \mathbb{C}^{3 \times 3}$ n'a qu'une seule valeur propre :

$$A = \begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & a \end{pmatrix} \Rightarrow \begin{aligned} \chi_A &= (X - a)^3 \\ \mu_A &= (X - a) \end{aligned}$$

$$A = \begin{pmatrix} a & 1 & 0 \\ 0 & a & 0 \\ 0 & 0 & a \end{pmatrix} \Rightarrow \begin{aligned} \chi_A &= (X - a)^3 \\ \mu_A &= (X - a)^2 \end{aligned}$$

$$A = \begin{pmatrix} a & 1 & 0 \\ 0 & a & 1 \\ 0 & 0 & a \end{pmatrix} \Rightarrow \chi_A = \mu_A = (X - a)^3$$

Corollaire

*Le polynôme minimal μ_A divise le polynôme caractéristique χ_A .
Il a les mêmes racines, éventuellement de multiplicité réduite.*

Comment calculer le polynôme minimal ?

On considère les matrices $I = A^0, A = A^1, A^2, \dots, A^n$ comme vecteurs dans $\mathbb{K}^{(n^2)}$.

Comment calculer le polynôme minimal ?

On considère les matrices $I = A^0, A = A^1, A^2, \dots, A^n$ comme vecteurs dans $\mathbb{K}^{(n^2)}$. On peut les écrire comme vecteurs lignes : Ceci donne une matrice M à $n + 1$ lignes et n^2 colonnes.

Comment calculer le polynôme minimal ?

On considère les matrices $I = A^0, A = A^1, A^2, \dots, A^n$ comme vecteurs dans $\mathbb{K}^{(n^2)}$. On peut les écrire comme vecteurs lignes : Ceci donne une matrice M à $n + 1$ lignes et n^2 colonnes.

Si les k premières lignes de M sont linéairement dépendantes, on obtient alors on relation non triviale

$$a_0 I + a_1 A + a_2 A^2 + \dots + a_k A^k = 0.$$

Comment calculer le polynôme minimal ?

On considère les matrices $I = A^0, A = A^1, A^2, \dots, A^n$ comme vecteurs dans $\mathbb{K}^{(n^2)}$. On peut les écrire comme vecteurs lignes : Ceci donne une matrice M à $n + 1$ lignes et n^2 colonnes.

Si les k premières lignes de M sont linéairement dépendantes, on obtient alors on relation non triviale

$$a_0 I + a_1 A + a_2 A^2 + \dots + a_k A^k = 0.$$

Ceci se détermine en appliquant l'algorithme de Gauss.

Comment calculer le polynôme minimal ?

On considère les matrices $I = A^0, A = A^1, A^2, \dots, A^n$ comme vecteurs dans $\mathbb{K}^{(n^2)}$. On peut les écrire comme vecteurs lignes : Ceci donne une matrice M à $n + 1$ lignes et n^2 colonnes.

Si les k premières lignes de M sont linéairement dépendantes, on obtient alors on relation non triviale

$$a_0I + a_1A + a_2A^2 + \dots + a_kA^k = 0.$$

Ceci se détermine en appliquant l'algorithme de Gauss.

Toute telle relation donne un polynôme annulateur

$$P = a_0 + a_1X + a_2X^2 + \dots + a_kX^k.$$

Comment calculer le polynôme minimal ?

On considère les matrices $I = A^0, A = A^1, A^2, \dots, A^n$ comme vecteurs dans $\mathbb{K}^{(n^2)}$. On peut les écrire comme vecteurs lignes : Ceci donne une matrice M à $n + 1$ lignes et n^2 colonnes.

Si les k premières lignes de M sont linéairement dépendantes, on obtient alors on relation non triviale

$$a_0 I + a_1 A + a_2 A^2 + \dots + a_k A^k = 0.$$

Ceci se détermine en appliquant l'algorithme de Gauss.

Toute telle relation donne un polynôme annulateur

$$P = a_0 + a_1 X + a_2 X^2 + \dots + a_k X^k.$$

Si k est minimal, alors $a_k \neq 0$ et on obtient $\mu_A = P/a_k$.

Un algorithme probabiliste plus rapide

Pour calculer μ_A il existe une méthode probabiliste moins coûteuse.

Un algorithme probabiliste plus rapide

Pour calculer μ_A il existe une méthode probabiliste moins coûteuse.
On calcule le polynôme minimal μ_A^v par rapport à un vecteur $v \neq 0$:

Un algorithme probabiliste plus rapide

Pour calculer μ_A il existe une méthode probabiliste moins coûteuse.

On calcule le polynôme minimal μ_A^v par rapport à un vecteur $v \neq 0$:

- On calcule les vecteurs $v_0 = v, v_1 = Av, \dots, v_n = A^n v$.

Un algorithme probabiliste plus rapide

Pour calculer μ_A il existe une méthode probabiliste moins coûteuse. On calcule le polynôme minimal μ_A^v par rapport à un vecteur $v \neq 0$:

- On calcule les vecteurs $v_0 = v, v_1 = Av, \dots, v_n = A^n v$.
- Dans \mathbb{K}^n on cherche une relation non triviale

$$a_0 v_0 + a_1 v_1 + \dots + a_k v_k = 0$$

avec k minimal. On peut ensuite se ramener à $a_k = 1$.

Un algorithme probabiliste plus rapide

Pour calculer μ_A il existe une méthode probabiliste moins coûteuse. On calcule le polynôme minimal μ_A^v par rapport à un vecteur $v \neq 0$:

- On calcule les vecteurs $v_0 = v, v_1 = Av, \dots, v_n = A^n v$.
- Dans \mathbb{K}^n on cherche une relation non triviale

$$a_0 v_0 + a_1 v_1 + \dots + a_k v_k = 0$$

avec k minimal. On peut ensuite se ramener à $a_k = 1$.

- Ceci définit un polynôme $\mu_A^v = a_0 + a_1 X + \dots + a_k X^k$.
Il vérifie $\mu_A^v(A)v = 0$, mais pas forcément $\mu_A^v(A) = 0$.

Un algorithme probabiliste plus rapide

Pour calculer μ_A il existe une méthode probabiliste moins coûteuse. On calcule le polynôme minimal μ_A^v par rapport à un vecteur $v \neq 0$:

- On calcule les vecteurs $v_0 = v, v_1 = Av, \dots, v_n = A^n v$.
- Dans \mathbb{K}^n on cherche une relation non triviale

$$a_0 v_0 + a_1 v_1 + \dots + a_k v_k = 0$$

avec k minimal. On peut ensuite se ramener à $a_k = 1$.

- Ceci définit un polynôme $\mu_A^v = a_0 + a_1 X + \dots + a_k X^k$.
Il vérifie $\mu_A^v(A)v = 0$, mais pas forcément $\mu_A^v(A) = 0$.

Avantage : les calculs sont beaucoup moins lourds !

La construction garantit que μ_A^v divise μ_A , qui lui divise χ_A .

Un algorithme probabiliste plus rapide

Pour calculer μ_A il existe une méthode probabiliste moins coûteuse. On calcule le polynôme minimal μ_A^v par rapport à un vecteur $v \neq 0$:

- On calcule les vecteurs $v_0 = v, v_1 = Av, \dots, v_n = A^n v$.
- Dans \mathbb{K}^n on cherche une relation non triviale

$$a_0 v_0 + a_1 v_1 + \dots + a_k v_k = 0$$

avec k minimal. On peut ensuite se ramener à $a_k = 1$.

- Ceci définit un polynôme $\mu_A^v = a_0 + a_1 X + \dots + a_k X^k$.
Il vérifie $\mu_A^v(A)v = 0$, mais pas forcément $\mu_A^v(A) = 0$.

Avantage : les calculs sont beaucoup moins lourds !

La construction garantit que μ_A^v divise μ_A , qui lui divise χ_A .

On effectue ce calcul pour un vecteur v choisi au hasard.

Si, par chance, μ_A^v est de degré n , alors $\mu_A^v = \mu_A = \chi_A$.

Un algorithme probabiliste plus rapide

Pour calculer μ_A il existe une méthode probabiliste moins coûteuse. On calcule le polynôme minimal μ_A^v par rapport à un vecteur $v \neq 0$:

- On calcule les vecteurs $v_0 = v, v_1 = Av, \dots, v_n = A^n v$.
- Dans \mathbb{K}^n on cherche une relation non triviale

$$a_0 v_0 + a_1 v_1 + \dots + a_k v_k = 0$$

avec k minimal. On peut ensuite se ramener à $a_k = 1$.

- Ceci définit un polynôme $\mu_A^v = a_0 + a_1 X + \dots + a_k X^k$.
Il vérifie $\mu_A^v(A)v = 0$, mais pas forcément $\mu_A^v(A) = 0$.

Avantage : les calculs sont beaucoup moins lourds !

La construction garantit que μ_A^v divise μ_A , qui lui divise χ_A .

On effectue ce calcul pour un vecteur v choisi au hasard.

Si, par chance, μ_A^v est de degré n , alors $\mu_A^v = \mu_A = \chi_A$.

Sinon, on peut essayer quelques autres vecteurs,

puis calculer le ppcm des polynômes ainsi obtenus.

Si l'on obtient un polynôme de degré n on conclut comme avant.

Un algorithme probabiliste plus rapide

Pour calculer μ_A il existe une méthode probabiliste moins coûteuse. On calcule le polynôme minimal μ_A^v par rapport à un vecteur $v \neq 0$:

- On calcule les vecteurs $v_0 = v, v_1 = Av, \dots, v_n = A^n v$.
- Dans \mathbb{K}^n on cherche une relation non triviale

$$a_0 v_0 + a_1 v_1 + \dots + a_k v_k = 0$$

avec k minimal. On peut ensuite se ramener à $a_k = 1$.

- Ceci définit un polynôme $\mu_A^v = a_0 + a_1 X + \dots + a_k X^k$.
Il vérifie $\mu_A^v(A)v = 0$, mais pas forcément $\mu_A^v(A) = 0$.

Avantage : les calculs sont beaucoup moins lourds !

La construction garantit que μ_A^v divise μ_A , qui lui divise χ_A .

On effectue ce calcul pour un vecteur v choisi au hasard.

Si, par chance, μ_A^v est de degré n , alors $\mu_A^v = \mu_A = \chi_A$.

Sinon, on peut essayer quelques autres vecteurs, puis calculer le ppcm des polynômes ainsi obtenus.

Si l'on obtient un polynôme de degré n on conclut comme avant.

Sinon on peut tester si ce polynôme annule la matrice A .

Dans ce cas favorable on a trouvé le polynôme minimal μ_A .

Diagonalisation de matrices : critères

On considère une matrice $A \in \mathbb{K}^{n \times n}$,
typiquement à coefficients dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$.

Diagonalisation de matrices : critères

On considère une matrice $A \in \mathbb{K}^{n \times n}$,
typiquement à coefficients dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$.

A est diagonalisable

Diagonalisation de matrices : critères

On considère une matrice $A \in \mathbb{K}^{n \times n}$,
typiquement à coefficients dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$.

A est *diagonalisable*

\Leftrightarrow Déf Il existe une base $v_1, \dots, v_n \in \mathbb{K}^n$ formée de vecteurs propres,
c'est-à-dire $Av_1 = \lambda_1 v_1, \dots, Av_n = \lambda_n v_n$.

Diagonalisation de matrices : critères

On considère une matrice $A \in \mathbb{K}^{n \times n}$,
typiquement à coefficients dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$.

A est *diagonalisable*

\Leftrightarrow ^{Déf} Il existe une base $v_1, \dots, v_n \in \mathbb{K}^n$ formée de vecteurs propres, c'est-à-dire $Av_1 = \lambda_1 v_1, \dots, Av_n = \lambda_n v_n$.

\Leftrightarrow Dans la base (v_1, \dots, v_n) l'endomorphisme $x \mapsto Ax$ s'écrit comme une matrice diagonale, $TAT^{-1} = \text{diag}(\lambda_1, \dots, \lambda_n)$. Ici T est la matrice de passage de (e_1, \dots, e_n) à (v_1, \dots, v_n) .

Diagonalisation de matrices : critères

On considère une matrice $A \in \mathbb{K}^{n \times n}$,
typiquement à coefficients dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$.

A est *diagonalisable*

- \Leftrightarrow ^{Déf} Il existe une base $v_1, \dots, v_n \in \mathbb{K}^n$ formée de vecteurs propres, c'est-à-dire $Av_1 = \lambda_1 v_1, \dots, Av_n = \lambda_n v_n$.
- \Leftrightarrow Dans la base (v_1, \dots, v_n) l'endomorphisme $x \mapsto Ax$ s'écrit comme une matrice diagonale, $TAT^{-1} = \text{diag}(\lambda_1, \dots, \lambda_n)$. Ici T est la matrice de passage de (e_1, \dots, e_n) à (v_1, \dots, v_n) .
- \Leftrightarrow Le polynôme minimal de A n'a que des racines simples, c'est-à-dire que $\text{pgcd}(\mu_A, \mu'_A) = 1$.

Diagonalisation de matrices : critères

On considère une matrice $A \in \mathbb{K}^{n \times n}$,
typiquement à coefficients dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$.

A est *diagonalisable*

- $\stackrel{\text{Déf}}{\iff}$ Il existe une base $v_1, \dots, v_n \in \mathbb{K}^n$ formée de vecteurs propres, c'est-à-dire $Av_1 = \lambda_1 v_1, \dots, Av_n = \lambda_n v_n$.
- \iff Dans la base (v_1, \dots, v_n) l'endomorphisme $x \mapsto Ax$ s'écrit comme une matrice diagonale, $TAT^{-1} = \text{diag}(\lambda_1, \dots, \lambda_n)$. Ici T est la matrice de passage de (e_1, \dots, e_n) à (v_1, \dots, v_n) .
- \iff Le polynôme minimal de A n'a que des racines simples, c'est-à-dire que $\text{pgcd}(\mu_A, \mu'_A) = 1$.
- \iff Le polynôme caractéristique de A n'a que des racines simples, c'est-à-dire que $\text{pgcd}(\chi_A, \chi'_A) = 1$.

Sommaire

- 1 Résolution de systèmes d'équations linéaires
- 2 Réduction des endomorphismes
- 3 Méthodes approchées itératives**
 - La méthode de la puissance
 - La méthode des itérations inverses
 - Matrices hermitiennes et symétriques
- 4 Comment fonctionne Google ?

La méthode de la puissance : idée

Considérons une matrice $A \in \mathbb{C}^{n \times n}$ à coefficients complexes.

La méthode de la puissance : idée

Considérons une matrice $A \in \mathbb{C}^{n \times n}$ à coefficients complexes.
Alors A admet n valeurs propres $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{C}$.

La méthode de la puissance : idée

Considérons une matrice $A \in \mathbb{C}^{n \times n}$ à coefficients complexes.

Alors A admet n valeurs propres $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{C}$.

Supposons que $|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|$.

La méthode de la puissance : idée

Considérons une matrice $A \in \mathbb{C}^{n \times n}$ à coefficients complexes.

Alors A admet n valeurs propres $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{C}$.

Supposons que $|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|$.

Objectif : Approcher un vecteur propre $v_1 \in \mathbb{C}^n$ associé à λ_1 .

La méthode de la puissance : idée

Considérons une matrice $A \in \mathbb{C}^{n \times n}$ à coefficients complexes.

Alors A admet n valeurs propres $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{C}$.

Supposons que $|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|$.

Objectif : Approcher un vecteur propre $v_1 \in \mathbb{C}^n$ associé à λ_1 .

Heuristique : Supposons que $A = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$.

La méthode de la puissance : idée

Considérons une matrice $A \in \mathbb{C}^{n \times n}$ à coefficients complexes.

Alors A admet n valeurs propres $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{C}$.

Supposons que $|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|$.

Objectif : Approcher un vecteur propre $v_1 \in \mathbb{C}^n$ associé à λ_1 .

Heuristique : Supposons que $A = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$.

Pour tout $x = (x_1, x_2, \dots, x_n) \in \mathbb{C}^n$ on trouve

$$Ax = (\lambda_1 x_1, \lambda_2 x_2, \dots, \lambda_n x_n),$$

$$A^k x = (\lambda_1^k x_1, \lambda_2^k x_2, \dots, \lambda_n^k x_n).$$

La méthode de la puissance : idée

Considérons une matrice $A \in \mathbb{C}^{n \times n}$ à coefficients complexes.

Alors A admet n valeurs propres $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{C}$.

Supposons que $|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|$.

Objectif : Approcher un vecteur propre $v_1 \in \mathbb{C}^n$ associé à λ_1 .

Heuristique : Supposons que $A = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$.

Pour tout $x = (x_1, x_2, \dots, x_n) \in \mathbb{C}^n$ on trouve

$$Ax = (\lambda_1 x_1, \lambda_2 x_2, \dots, \lambda_n x_n),$$

$$A^k x = (\lambda_1^k x_1, \lambda_2^k x_2, \dots, \lambda_n^k x_n).$$

La première coordonnée croît le plus vite ! (On suppose $x_1 \neq 0$.)

La méthode de la puissance : idée

Considérons une matrice $A \in \mathbb{C}^{n \times n}$ à coefficients complexes.

Alors A admet n valeurs propres $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{C}$.

Supposons que $|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|$.

Objectif : Approcher un vecteur propre $v_1 \in \mathbb{C}^n$ associé à λ_1 .

Heuristique : Supposons que $A = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$.

Pour tout $x = (x_1, x_2, \dots, x_n) \in \mathbb{C}^n$ on trouve

$$\begin{aligned}Ax &= (\lambda_1 x_1, \lambda_2 x_2, \dots, \lambda_n x_n), \\A^k x &= (\lambda_1^k x_1, \lambda_2^k x_2, \dots, \lambda_n^k x_n).\end{aligned}$$

La première coordonnée croît le plus vite ! (On suppose $x_1 \neq 0$.)

Méthode : On commence par un vecteur aléatoire $u_0 \in \mathbb{C}^n$ non nul.

La méthode de la puissance : idée

Considérons une matrice $A \in \mathbb{C}^{n \times n}$ à coefficients complexes.

Alors A admet n valeurs propres $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{C}$.

Supposons que $|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|$.

Objectif : Approcher un vecteur propre $v_1 \in \mathbb{C}^n$ associé à λ_1 .

Heuristique : Supposons que $A = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$.

Pour tout $x = (x_1, x_2, \dots, x_n) \in \mathbb{C}^n$ on trouve

$$\begin{aligned}Ax &= (\lambda_1 x_1, \lambda_2 x_2, \dots, \lambda_n x_n), \\A^k x &= (\lambda_1^k x_1, \lambda_2^k x_2, \dots, \lambda_n^k x_n).\end{aligned}$$

La première coordonnée croît le plus vite ! (On suppose $x_1 \neq 0$.)

Méthode : On commence par un vecteur aléatoire $u_0 \in \mathbb{C}^n$ non nul.

De manière itérative on calcule $u_{k+1} = \frac{Au_k}{|Au_k|}$ pour $k = 0, 1, 2, \dots$

La méthode de la puissance : idée

Considérons une matrice $A \in \mathbb{C}^{n \times n}$ à coefficients complexes.

Alors A admet n valeurs propres $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{C}$.

Supposons que $|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|$.

Objectif : Approcher un vecteur propre $v_1 \in \mathbb{C}^n$ associé à λ_1 .

Heuristique : Supposons que $A = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$.

Pour tout $x = (x_1, x_2, \dots, x_n) \in \mathbb{C}^n$ on trouve

$$\begin{aligned}Ax &= (\lambda_1 x_1, \lambda_2 x_2, \dots, \lambda_n x_n), \\ A^k x &= (\lambda_1^k x_1, \lambda_2^k x_2, \dots, \lambda_n^k x_n).\end{aligned}$$

La première coordonnée croît le plus vite ! (On suppose $x_1 \neq 0$.)

Méthode : On commence par un vecteur aléatoire $u_0 \in \mathbb{C}^n$ non nul.

De manière itérative on calcule $u_{k+1} = \frac{Au_k}{|Au_k|}$ pour $k = 0, 1, 2, \dots$

On s'arrête si $Au_k \approx \lambda u_k$ en approximation suffisante pour un $\lambda \in \mathbb{C}$.

La méthode de la puissance : idée

Considérons une matrice $A \in \mathbb{C}^{n \times n}$ à coefficients complexes.

Alors A admet n valeurs propres $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{C}$.

Supposons que $|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|$.

Objectif : Approcher un vecteur propre $v_1 \in \mathbb{C}^n$ associé à λ_1 .

Heuristique : Supposons que $A = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$.

Pour tout $x = (x_1, x_2, \dots, x_n) \in \mathbb{C}^n$ on trouve

$$\begin{aligned}Ax &= (\lambda_1 x_1, \lambda_2 x_2, \dots, \lambda_n x_n), \\A^k x &= (\lambda_1^k x_1, \lambda_2^k x_2, \dots, \lambda_n^k x_n).\end{aligned}$$

La première coordonnée croît le plus vite ! (On suppose $x_1 \neq 0$.)

Méthode : On commence par un vecteur aléatoire $u_0 \in \mathbb{C}^n$ non nul.

De manière itérative on calcule $u_{k+1} = \frac{Au_k}{|Au_k|}$ pour $k = 0, 1, 2, \dots$

On s'arrête si $Au_k \approx \lambda u_k$ en approximation suffisante pour un $\lambda \in \mathbb{C}$.

Dans ce cas $u_k \approx v_1$ et $\lambda \approx \lambda_1$ répondent au problème.

La méthode de la puissance : idée

Considérons une matrice $A \in \mathbb{C}^{n \times n}$ à coefficients complexes.

Alors A admet n valeurs propres $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{C}$.

Supposons que $|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|$.

Objectif : Approcher un vecteur propre $v_1 \in \mathbb{C}^n$ associé à λ_1 .

Heuristique : Supposons que $A = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$.

Pour tout $x = (x_1, x_2, \dots, x_n) \in \mathbb{C}^n$ on trouve

$$\begin{aligned}Ax &= (\lambda_1 x_1, \lambda_2 x_2, \dots, \lambda_n x_n), \\A^k x &= (\lambda_1^k x_1, \lambda_2^k x_2, \dots, \lambda_n^k x_n).\end{aligned}$$

La première coordonnée croît le plus vite ! (On suppose $x_1 \neq 0$.)

Méthode : On commence par un vecteur aléatoire $u_0 \in \mathbb{C}^n$ non nul.

De manière itérative on calcule $u_{k+1} = \frac{Au_k}{|Au_k|}$ pour $k = 0, 1, 2, \dots$

On s'arrête si $Au_k \approx \lambda u_k$ en approximation suffisante pour un $\lambda \in \mathbb{C}$.

Dans ce cas $u_k \approx v_1$ et $\lambda \approx \lambda_1$ répondent au problème.

Vitesse de convergence : L'erreur diminue comme k^n où $k = \frac{|\lambda_2|}{|\lambda_1|}$.

Plus λ_1 dépasse les autres valeurs $\lambda_2, \dots, \lambda_n$, mieux c'est !

Sommaire

- 1 Résolution de systèmes d'équations linéaires
- 2 Réduction des endomorphismes
- 3 Méthodes approchées itératives
- 4 Comment fonctionne Google ?**
 - Comment mesurer l'importance d'une page web ?
 - Le modèle PageRank : marche aléatoire sur le web
 - Existence, unicité, et calcul de la solution

Que fait un moteur de recherche ?

Que fait un moteur de recherche ?

Que fait un moteur de recherche ?

Que fait un moteur de recherche ?

- Fouille de données : pour un ou plusieurs mots-clés donnés on cherche des pages web associées, si possible pertinentes.

Que fait un moteur de recherche ?

Que fait un moteur de recherche ?

- Fouille de données : pour un ou plusieurs mots-clés donnés on cherche des pages web associées, si possible pertinentes.
- Classement des résultats : puisque les pages trouvées sont (souvent trop) nombreuses, il faut les trier par importance.

Que fait un moteur de recherche ?

Que fait un moteur de recherche ?

- Fouille de données : pour un ou plusieurs mots-clés donnés on cherche des pages web associées, si possible pertinentes.
- Classement des résultats : puisque les pages trouvées sont (souvent trop) nombreuses, il faut les trier par importance.

Derrière les coulisses, plusieurs problèmes se présentent :

Que fait un moteur de recherche ?

Que fait un moteur de recherche ?

- Fouille de données : pour un ou plusieurs mots-clés donnés on cherche des pages web associées, si possible pertinentes.
- Classement des résultats : puisque les pages trouvées sont (souvent trop) nombreuses, il faut les trier par importance.

Derrière les coulisses, plusieurs problèmes se présentent :

- La quantité des données à gérer est énorme.

Que fait un moteur de recherche ?

Que fait un moteur de recherche ?

- Fouille de données : pour un ou plusieurs mots-clés donnés on cherche des pages web associées, si possible pertinentes.
- Classement des résultats : puisque les pages trouvées sont (souvent trop) nombreuses, il faut les trier par importance.

Derrière les coulisses, plusieurs problèmes se présentent :

- La quantité des données à gérer est énorme.
- Toute requête doit être traitée en temps réel.

Que fait un moteur de recherche ?

Que fait un moteur de recherche ?

- Fouille de données : pour un ou plusieurs mots-clés donnés on cherche des pages web associées, si possible pertinentes.
- Classement des résultats : puisque les pages trouvées sont (souvent trop) nombreuses, il faut les trier par importance.

Derrière les coulisses, plusieurs problèmes se présentent :

- La quantité des données à gérer est énorme.
- Toute requête doit être traitée en temps réel.

Finalement, il y a un problème encore plus délicat :

Que fait un moteur de recherche ?

Que fait un moteur de recherche ?

- Fouille de données : pour un ou plusieurs mots-clés donnés on cherche des pages web associées, si possible pertinentes.
- Classement des résultats : puisque les pages trouvées sont (souvent trop) nombreuses, il faut les trier par importance.

Derrière les coulisses, plusieurs problèmes se présentent :

- La quantité des données à gérer est énorme.
- Toute requête doit être traitée en temps réel.

Finalement, il y a un problème encore plus délicat :

- Comment trier les pages par ordre d'importance ?

Le web est un graphe

Dans une première approximation nous allons négliger le contenu des pages et ne tenir compte que de la structure du graphe.

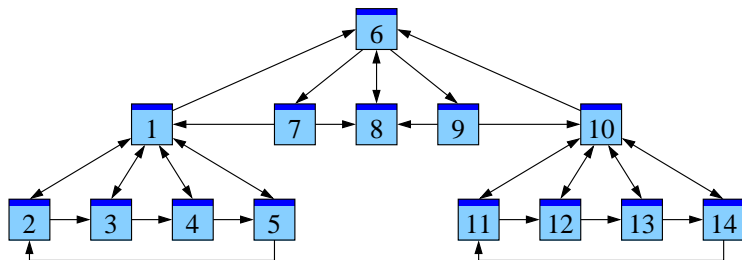


FIG.: Le web vu comme un graphe

Le web est un graphe

Dans une première approximation nous allons négliger le contenu des pages et ne tenir compte que de la structure du graphe.

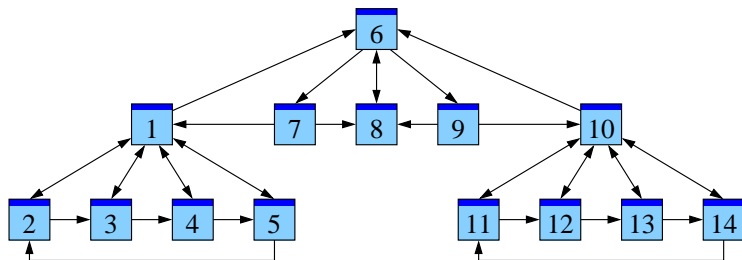


FIG.: Le web vu comme un graphe

Dans la suite on numérote les pages par $1, 2, 3, \dots, n$.

Le web est un graphe

Dans une première approximation nous allons négliger le contenu des pages et ne tenir compte que de la structure du graphe.

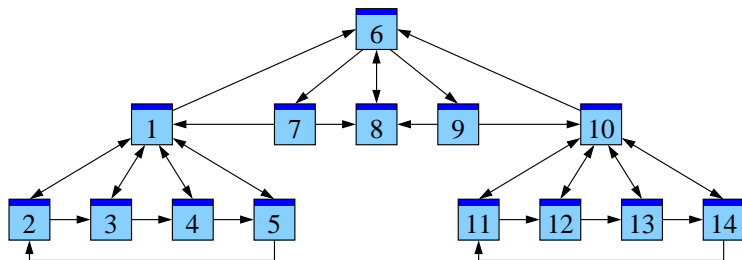


FIG.: Le web vu comme un graphe

Dans la suite on numérote les pages par $1, 2, 3, \dots, n$.
On écrit $j \rightarrow i$ si la page j pointe vers la page i .

Le web est un graphe

Dans une première approximation nous allons négliger le contenu des pages et ne tenir compte que de la structure du graphe.

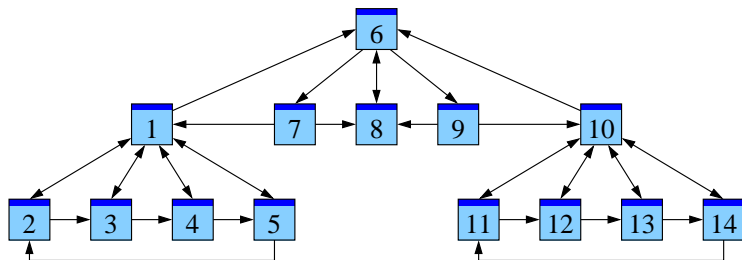


FIG.: Le web vu comme un graphe

Dans la suite on numérote les pages par $1, 2, 3, \dots, n$.

On écrit $j \rightarrow i$ si la page j pointe vers la page i . Exemple :

$1 \rightarrow 2, 3, 4, 5, 6$; $2 \rightarrow 1, 3$; $3 \rightarrow 1, 4$; $4 \rightarrow 1, 5$; $5 \rightarrow 1, 2$;

$6 \rightarrow 7, 8, 9$; $7 \rightarrow 8, 1$; $8 \rightarrow 6$; $9 \rightarrow 8, 10$; $10 \rightarrow 6, 11, 12, 13, 14$;

$11 \rightarrow 10, 12$; $12 \rightarrow 10, 13$; $13 \rightarrow 10, 14$; $14 \rightarrow 10, 11$.

Comment mesurer l'importance d'une page web ?

Première idée : comptage des liens.

Il est plausible qu'une page importante reçoit beaucoup de liens.

Comment mesurer l'importance d'une page web ?

Première idée : comptage des liens.

Il est plausible qu'une page importante reçoit beaucoup de liens.

Avec un peu de naïveté, on croira aussi l'affirmation réciproque : si une page reçoit beaucoup de liens, alors elle est importante.

Comment mesurer l'importance d'une page web ?

Première idée : comptage des liens.

Il est plausible qu'une page importante reçoit beaucoup de liens.

Avec un peu de naïveté, on croira aussi l'affirmation réciproque : si une page reçoit beaucoup de liens, alors elle est importante.

Ainsi on pourrait définir l'importance μ_i de la page i comme suit :

$$(1) \quad \mu_i = \sum_{j \rightarrow i} 1.$$

Comment mesurer l'importance d'une page web ?

Première idée : comptage des liens.

Il est plausible qu'une page importante reçoit beaucoup de liens.

Avec un peu de naïveté, on croira aussi l'affirmation réciproque : si une page reçoit beaucoup de liens, alors elle est importante.

Ainsi on pourrait définir l'importance μ_i de la page i comme suit :

$$(1) \quad \mu_i = \sum_{j \rightarrow i} 1.$$

Exemple : Les pages 1 et 10 reçoivent 5 liens chacune, alors que la page 6 n'en reçoit que 3. Ainsi $\mu_1 = \mu_{10} = 5$ mais seulement $\mu_6 = 3$.

Comment mesurer l'importance d'une page web ?

Première idée : comptage des liens.

Il est plausible qu'une page importante reçoit beaucoup de liens.

Avec un peu de naïveté, on croira aussi l'affirmation réciproque : si une page reçoit beaucoup de liens, alors elle est importante.

Ainsi on pourrait définir l'importance μ_i de la page i comme suit :

$$(1) \quad \mu_i = \sum_{j \rightarrow i} 1.$$

Exemple : Les pages 1 et 10 reçoivent 5 liens chacune, alors que la page 6 n'en reçoit que 3. Ainsi $\mu_1 = \mu_{10} = 5$ mais seulement $\mu_6 = 3$.

Inconvénient : La mesure μ ainsi définie ne correspond pas à l'importance ressentie : elle sous-estime l'importance de la page 6.

Comment mesurer l'importance d'une page web ?

Première idée : comptage des liens.

Il est plausible qu'une page importante reçoit beaucoup de liens.

Avec un peu de naïveté, on croira aussi l'affirmation réciproque : si une page reçoit beaucoup de liens, alors elle est importante.

Ainsi on pourrait définir l'importance μ_i de la page i comme suit :

$$(1) \quad \mu_i = \sum_{j \rightarrow i} 1.$$

Exemple : Les pages 1 et 10 reçoivent 5 liens chacune, alors que la page 6 n'en reçoit que 3. Ainsi $\mu_1 = \mu_{10} = 5$ mais seulement $\mu_6 = 3$.

Inconvénient : La mesure μ ainsi définie ne correspond pas à l'importance ressentie : elle sous-estime l'importance de la page 6.

Manipulation : On peut artificiellement augmenter l'importance d'une page i en créant des pages « vides de sens » pointant vers i .

Comment mesurer l'importance d'une page web ?

Seconde idée : comptage pondéré.

Certaines pages j émettent un grand nombre ℓ_j de liens :
ceux-ci sont donc moins spécifiques et leur « poids » est plus faible.

Comment mesurer l'importance d'une page web ?

Seconde idée : comptage pondéré.

Certaines pages j émettent un grand nombre ℓ_j de liens : ceux-ci sont donc moins spécifiques et leur « poids » est plus faible. Ainsi on pourrait définir une mesure d'importance plus fine par

$$(2) \quad \mu_i = \sum_{j \rightarrow i} \frac{1}{\ell_j}.$$

Comment mesurer l'importance d'une page web ?

Seconde idée : comptage pondéré.

Certaines pages j émettent un grand nombre ℓ_j de liens : ceux-ci sont donc moins spécifiques et leur « poids » est plus faible. Ainsi on pourrait définir une mesure d'importance plus fine par

$$(2) \quad \mu_i = \sum_{j \rightarrow i} \frac{1}{\ell_j}.$$

Exemple : Dans notre exemple, on trouve $\mu_1 = \mu_{10} = 2.5$ et $\mu_6 = 1.4$.

Comment mesurer l'importance d'une page web ?

Seconde idée : comptage pondéré.

Certaines pages j émettent un grand nombre ℓ_j de liens : ceux-ci sont donc moins spécifiques et leur « poids » est plus faible. Ainsi on pourrait définir une mesure d'importance plus fine par

$$(2) \quad \mu_i = \sum_{j \rightarrow i} \frac{1}{\ell_j}.$$

Exemple : Dans notre exemple, on trouve $\mu_1 = \mu_{10} = 2.5$ et $\mu_6 = 1.4$.

Inconvénient : La mesure μ ainsi définie ne correspond toujours pas bien à l'importance ressentie par les utilisateurs.

Comment mesurer l'importance d'une page web ?

Seconde idée : comptage pondéré.

Certaines pages j émettent un grand nombre ℓ_j de liens : ceux-ci sont donc moins spécifiques et leur « poids » est plus faible. Ainsi on pourrait définir une mesure d'importance plus fine par

$$(2) \quad \mu_i = \sum_{j \rightarrow i} \frac{1}{\ell_j}.$$

Exemple : Dans notre exemple, on trouve $\mu_1 = \mu_{10} = 2.5$ et $\mu_6 = 1.4$.

Inconvénient : La mesure μ ainsi définie ne correspond toujours pas bien à l'importance ressentie par les utilisateurs.

Manipulation : Comme avant, on peut artificiellement augmenter l'importance d'une page i en créant une foule de pages « vides » pointant vers i . De nouveau, la mesure n'est pas fiable.

Comment mesurer l'importance d'une page web ?

Troisième idée : définition récursive.

Le principe utilisé par Google : *une page i est importante si beaucoup de pages importantes pointent vers i .*

Comment mesurer l'importance d'une page web ?

Troisième idée : définition récursive.

Le principe utilisé par Google : *une page i est importante si beaucoup de pages importantes pointent vers i .*

Ainsi on pose

$$(3) \quad \mu_i = \sum_{j \rightarrow i} \frac{1}{\ell_j} \mu_j.$$

Comment mesurer l'importance d'une page web ?

Troisième idée : définition réursive.

Le principe utilisé par Google : *une page i est importante si beaucoup de pages importantes pointent vers i .*

Ainsi on pose

$$(3) \quad \mu_i = \sum_{j \rightarrow i} \frac{1}{\ell_j} \mu_j.$$

Exemple : Dans notre exemple, on trouve $\mu_6 = 6$ et $\mu_1 = \mu_{10} = 5$ puis $\mu_8 = 4$. Les autres pages n'obtiennent que $\mu_i = 2$.

Comment mesurer l'importance d'une page web ?

Troisième idée : définition récursive.

Le principe utilisé par Google : *une page i est importante si beaucoup de pages importantes pointent vers i .*

Ainsi on pose

$$(3) \quad \mu_i = \sum_{j \rightarrow i} \frac{1}{\ell_j} \mu_j.$$

Exemple : Dans notre exemple, on trouve $\mu_6 = 6$ et $\mu_1 = \mu_{10} = 5$ puis $\mu_8 = 4$. Les autres pages n'obtiennent que $\mu_i = 2$.

Plausibilité : Les pages 6, 1, 10, 8 sont effectivement repérées comme les plus importantes.

Comment mesurer l'importance d'une page web ?

Troisième idée : définition récursive.

Le principe utilisé par Google : *une page i est importante si beaucoup de pages importantes pointent vers i .*

Ainsi on pose

$$(3) \quad \mu_i = \sum_{j \rightarrow i} \frac{1}{\ell_j} \mu_j.$$

Exemple : Dans notre exemple, on trouve $\mu_6 = 6$ et $\mu_1 = \mu_{10} = 5$ puis $\mu_8 = 4$. Les autres pages n'obtiennent que $\mu_i = 2$.

Plausibilité : Les pages 6, 1, 10, 8 sont effectivement repérées comme les plus importantes.

Robustesse : Si l'on ajoute des pages « vides de sens » elles recevront l'importance 0 et ne contribueront pas au calcul.

Interprétation en termes d'algèbre linéaire

Notre formule n'est rien autre qu'un système d'équations linéaires :

$$(4) \quad \mu_i = \sum_{j \rightarrow i} \frac{1}{\ell_j} \mu_j.$$

Interprétation en termes d'algèbre linéaire

Notre formule n'est rien autre qu'un système d'équations linéaires :

$$(4) \quad \mu_i = \sum_{j \rightarrow i} \frac{1}{\ell_j} \mu_j.$$

Plus explicitement, pour $i, j \in \{1, \dots, n\}$, on pose

$$(5) \quad a_{ij} := \begin{cases} \frac{1}{\ell_j} & \text{si } j \rightarrow i, \\ 0 & \text{sinon.} \end{cases}$$

Interprétation en termes d'algèbre linéaire

Notre formule n'est rien autre qu'un système d'équations linéaires :

$$(4) \quad \mu_i = \sum_{j \rightarrow i} \frac{1}{\ell_j} \mu_j.$$

Plus explicitement, pour $i, j \in \{1, \dots, n\}$, on pose

$$(5) \quad a_{ij} := \begin{cases} \frac{1}{\ell_j} & \text{si } j \rightarrow i, \\ 0 & \text{sinon.} \end{cases}$$

Avec cette matrice $A = (a_{ij})$ notre équation (4) s'écrit comme

$$(6) \quad \mu = A\mu$$

Interprétation en termes d'algèbre linéaire

Notre formule n'est rien autre qu'un système d'équations linéaires :

$$(4) \quad \mu_i = \sum_{j \rightarrow i} \frac{1}{\ell_j} \mu_j.$$

Plus explicitement, pour $i, j \in \{1, \dots, n\}$, on pose

$$(5) \quad a_{ij} := \begin{cases} \frac{1}{\ell_j} & \text{si } j \rightarrow i, \\ 0 & \text{sinon.} \end{cases}$$

Avec cette matrice $A = (a_{ij})$ notre équation (4) s'écrit comme

$$(6) \quad \mu = A\mu$$

C'est un système d'équations linéaires, que l'on peut résoudre par des méthodes adéquates. Remarquons qu'il se réécrit comme

$$(7) \quad (A - I)\mu = 0.$$

Interprétation en termes d'algèbre linéaire

Notre formule n'est rien autre qu'un système d'équations linéaires :

$$(4) \quad \mu_i = \sum_{j \rightarrow i} \frac{1}{\ell_j} \mu_j.$$

Plus explicitement, pour $i, j \in \{1, \dots, n\}$, on pose

$$(5) \quad a_{ij} := \begin{cases} \frac{1}{\ell_j} & \text{si } j \rightarrow i, \\ 0 & \text{sinon.} \end{cases}$$

Avec cette matrice $A = (a_{ij})$ notre équation (4) s'écrit comme

$$(6) \quad \mu = A\mu$$

C'est un système d'équations linéaires, que l'on peut résoudre par des méthodes adéquates. Remarquons qu'il se réécrit comme

$$(7) \quad (A - I)\mu = 0.$$

Nous cherchons un vecteur propre $\mu \in \mathbb{R}^n$ pour la valeur propre 1.

Interprétation comme marche aléatoire

Avant de calculer aveuglement, interprétons notre équation $A\mu = \mu$.

Interprétation comme marche aléatoire

Avant de calculer aveuglement, interprétons notre équation $A\mu = \mu$.

Par définition $A = (a_{ij})$ est une *matrice stochastique* :

$$\begin{array}{ll} a_{ij} \geq 0 & \text{pour tout } i, j \text{ et} \\ \sum_i a_{ij} = 1 & \text{pour tout } j. \end{array}$$

Interprétation comme marche aléatoire

Avant de calculer aveuglement, interprétons notre équation $A\mu = \mu$.

Par définition $A = (a_{ij})$ est une *matrice stochastique* :

$$\begin{aligned} a_{ij} &\geq 0 && \text{pour tout } i, j \text{ et} \\ \sum_i a_{ij} &= 1 && \text{pour tout } j. \end{aligned}$$

On interprète a_{ij} comme la probabilité d'aller de la page j à la page i en suivant un des ℓ_j liens, choisi au hasard.

Interprétation comme marche aléatoire

Avant de calculer aveuglement, interprétons notre équation $A\mu = \mu$.

Par définition $A = (a_{ij})$ est une *matrice stochastique* :

$$\begin{aligned} a_{ij} &\geq 0 && \text{pour tout } i, j \text{ et} \\ \sum_i a_{ij} &= 1 && \text{pour tout } j. \end{aligned}$$

On interprète a_{ij} comme la probabilité d'aller de la page j à la page i en suivant un des ℓ_j liens, choisi au hasard.

Supposons que $x \in \mathbb{R}^n$ est un *vecteur stochastique* :

$$x_j \geq 0 \quad \text{pour tout } j \text{ et } \quad \sum_j x_j = 1,$$

Interprétation comme marche aléatoire

Avant de calculer aveuglement, interprétons notre équation $A\mu = \mu$.

Par définition $A = (a_{ij})$ est une *matrice stochastique* :

$$\begin{aligned} a_{ij} &\geq 0 && \text{pour tout } i, j \text{ et} \\ \sum_i a_{ij} &= 1 && \text{pour tout } j. \end{aligned}$$

On interprète a_{ij} comme la probabilité d'aller de la page j à la page i en suivant un des ℓ_j liens, choisi au hasard.

Supposons que $x \in \mathbb{R}^n$ est un *vecteur stochastique* :

$$x_j \geq 0 \quad \text{pour tout } j \text{ et } \quad \sum_j x_j = 1,$$

On interprète x_j comme la probabilité de se trouver sur la page j .

Interprétation comme marche aléatoire

Avant de calculer aveuglement, interprétons notre équation $A\mu = \mu$.

Par définition $A = (a_{ij})$ est une *matrice stochastique* :

$$\begin{aligned} a_{ij} &\geq 0 && \text{pour tout } i, j \text{ et} \\ \sum_i a_{ij} &= 1 && \text{pour tout } j. \end{aligned}$$

On interprète a_{ij} comme la probabilité d'aller de la page j à la page i en suivant un des ℓ_j liens, choisi au hasard.

Supposons que $x \in \mathbb{R}^n$ est un *vecteur stochastique* :

$$x_j \geq 0 \quad \text{pour tout } j \text{ et } \quad \sum_j x_j = 1,$$

On interprète x_j comme la probabilité de se trouver sur la page j .

Partant de la page j , on suit le lien $j \rightarrow i$ avec probabilité a_{ij} .

Interprétation comme marche aléatoire

Avant de calculer aveuglement, interprétons notre équation $A\mu = \mu$.

Par définition $A = (a_{ij})$ est une *matrice stochastique* :

$$\begin{aligned} a_{ij} &\geq 0 && \text{pour tout } i, j \text{ et} \\ \sum_i a_{ij} &= 1 && \text{pour tout } j. \end{aligned}$$

On interprète a_{ij} comme la probabilité d'aller de la page j à la page i en suivant un des ℓ_j liens, choisi au hasard.

Supposons que $x \in \mathbb{R}^n$ est un *vecteur stochastique* :

$$x_j \geq 0 \quad \text{pour tout } j \text{ et } \quad \sum_j x_j = 1,$$

On interprète x_j comme la probabilité de se trouver sur la page j .

Partant de la page j , on suit le lien $j \rightarrow i$ avec probabilité a_{ij} .

Ce chemin nous fait tomber sur la page i avec une probabilité $a_{ij}x_j$.

Interprétation comme marche aléatoire

Avant de calculer aveuglement, interprétons notre équation $A\mu = \mu$.

Par définition $A = (a_{ij})$ est une *matrice stochastique* :

$$\begin{aligned} a_{ij} &\geq 0 && \text{pour tout } i, j \text{ et} \\ \sum_i a_{ij} &= 1 && \text{pour tout } j. \end{aligned}$$

On interprète a_{ij} comme la probabilité d'aller de la page j à la page i en suivant un des ℓ_j liens, choisi au hasard.

Supposons que $x \in \mathbb{R}^n$ est un *vecteur stochastique* :

$$x_j \geq 0 \quad \text{pour tout } j \text{ et } \quad \sum_j x_j = 1,$$

On interprète x_j comme la probabilité de se trouver sur la page j .

Partant de la page j , on suit le lien $j \rightarrow i$ avec probabilité a_{ij} .

Ce chemin nous fait tomber sur la page i avec une probabilité $a_{ij}x_j$.

La probabilité d'arriver sur la page i , par n'importe quel chemin, est

$$y_i = \sum_j a_{ij}x_j.$$

Interprétation comme marche aléatoire

Avant de calculer aveuglement, interprétons notre équation $A\mu = \mu$.

Par définition $A = (a_{ij})$ est une *matrice stochastique* :

$$\begin{aligned} a_{ij} &\geq 0 && \text{pour tout } i, j \text{ et} \\ \sum_i a_{ij} &= 1 && \text{pour tout } j. \end{aligned}$$

On interprète a_{ij} comme la probabilité d'aller de la page j à la page i en suivant un des ℓ_j liens, choisi au hasard.

Supposons que $x \in \mathbb{R}^n$ est un *vecteur stochastique* :

$$x_j \geq 0 \quad \text{pour tout } j \text{ et } \quad \sum_j x_j = 1,$$

On interprète x_j comme la probabilité de se trouver sur la page j .

Partant de la page j , on suit le lien $j \rightarrow i$ avec probabilité a_{ij} .

Ce chemin nous fait tomber sur la page i avec une probabilité $a_{ij}x_j$.

La probabilité d'arriver sur la page i , par n'importe quel chemin, est

$$y_i = \sum_j a_{ij}x_j.$$

Un pas dans la marche aléatoire correspond à l'application linéaire

$$T: \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad x \mapsto y = Ax.$$

Mesure invariante

Une mesure de probabilité μ vérifiant $\mu = T(\mu)$ est appelée une *mesure invariante* ou une *mesure d'équilibre*.

Mesure invariante

Une mesure de probabilité μ vérifiant $\mu = T(\mu)$ est appelée une *mesure invariante* ou une *mesure d'équilibre*.

En termes d'algèbre linéaire c'est un *vecteur propre* pour $\lambda = 1$.

Mesure invariante

Une mesure de probabilité μ vérifiant $\mu = T(\mu)$ est appelée une *mesure invariante* ou une *mesure d'équilibre*.

En termes d'algèbre linéaire c'est un *vecteur propre* pour $\lambda = 1$.

En termes d'analyse, c'est un *point fixe* de l'application T .

Mesure invariante

Une mesure de probabilité μ vérifiant $\mu = T(\mu)$ est appelée une *mesure invariante* ou une *mesure d'équilibre*.

En termes d'algèbre linéaire c'est un *vecteur propre* pour $\lambda = 1$.

En termes d'analyse, c'est un *point fixe* de l'application T .

Exemple : itérons la marche aléatoire avec une probabilité initiale u_0 :

temps	page 1	page 2	page 3	page 4	page 5	page 6	page 7	page 8	page 9	page 10	page 11	page 12	page 13	page 14
$t=0$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000
$t=1$	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$t=2$	0.000	0.000	0.000	0.000	0.000	0.000	0.333	0.333	0.333	0.000	0.000	0.000	0.000	0.000
$t=3$	0.167	0.000	0.000	0.000	0.000	0.333	0.000	0.333	0.000	0.167	0.000	0.000	0.000	0.000
$t=4$	0.000	0.033	0.033	0.033	0.033	0.400	0.111	0.111	0.111	0.000	0.033	0.033	0.033	0.033
$t=5$	0.122	0.017	0.017	0.017	0.017	0.111	0.133	0.244	0.133	0.122	0.017	0.017	0.017	0.017
$t=6$	0.100	0.033	0.033	0.033	0.033	0.293	0.037	0.170	0.037	0.100	0.033	0.033	0.033	0.033
$t=7$	0.084	0.036	0.036	0.036	0.036	0.210	0.098	0.135	0.098	0.084	0.036	0.036	0.036	0.036
$t=8$	0.122	0.035	0.035	0.035	0.035	0.168	0.070	0.168	0.070	0.122	0.035	0.035	0.035	0.035
$t=9$	0.105	0.042	0.042	0.042	0.042	0.217	0.056	0.126	0.056	0.105	0.042	0.042	0.042	0.042
...														
$t=28$	0.125	0.050	0.050	0.050	0.050	0.151	0.050	0.100	0.050	0.125	0.050	0.050	0.050	0.050
$t=29$	0.125	0.050	0.050	0.050	0.050	0.150	0.050	0.100	0.050	0.125	0.050	0.050	0.050	0.050
$t=30$	0.125	0.050	0.050	0.050	0.050	0.150	0.050	0.100	0.050	0.125	0.050	0.050	0.050	0.050

Mesure invariante

Une mesure de probabilité μ vérifiant $\mu = T(\mu)$ est appelée une *mesure invariante* ou une *mesure d'équilibre*.

En termes d'algèbre linéaire c'est un *vecteur propre* pour $\lambda = 1$.

En termes d'analyse, c'est un *point fixe* de l'application T .

Exemple : itérons la marche aléatoire avec une probabilité initiale u_0 :

temps	page 1	page 2	page 3	page 4	page 5	page 6	page 7	page 8	page 9	page 10	page 11	page 12	page 13	page 14
$t=0$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000
$t=1$	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$t=2$	0.000	0.000	0.000	0.000	0.000	0.000	0.333	0.333	0.333	0.000	0.000	0.000	0.000	0.000
$t=3$	0.167	0.000	0.000	0.000	0.000	0.333	0.000	0.333	0.000	0.167	0.000	0.000	0.000	0.000
$t=4$	0.000	0.033	0.033	0.033	0.033	0.400	0.111	0.111	0.111	0.000	0.033	0.033	0.033	0.033
$t=5$	0.122	0.017	0.017	0.017	0.017	0.111	0.133	0.244	0.133	0.122	0.017	0.017	0.017	0.017
$t=6$	0.100	0.033	0.033	0.033	0.033	0.293	0.037	0.170	0.037	0.100	0.033	0.033	0.033	0.033
$t=7$	0.084	0.036	0.036	0.036	0.036	0.210	0.098	0.135	0.098	0.084	0.036	0.036	0.036	0.036
$t=8$	0.122	0.035	0.035	0.035	0.035	0.168	0.070	0.168	0.070	0.122	0.035	0.035	0.035	0.035
$t=9$	0.105	0.042	0.042	0.042	0.042	0.217	0.056	0.126	0.056	0.105	0.042	0.042	0.042	0.042
...														
$t=28$	0.125	0.050	0.050	0.050	0.050	0.151	0.050	0.100	0.050	0.125	0.050	0.050	0.050	0.050
$t=29$	0.125	0.050	0.050	0.050	0.050	0.150	0.050	0.100	0.050	0.125	0.050	0.050	0.050	0.050
$t=30$	0.125	0.050	0.050	0.050	0.050	0.150	0.050	0.100	0.050	0.125	0.050	0.050	0.050	0.050

On observe un phénomène de diffusion, plausible après réflexion.

Mesure invariante

Une mesure de probabilité μ vérifiant $\mu = T(\mu)$ est appelée une *mesure invariante* ou une *mesure d'équilibre*.

En termes d'algèbre linéaire c'est un *vecteur propre* pour $\lambda = 1$.

En termes d'analyse, c'est un *point fixe* de l'application T .

Exemple : itérons la marche aléatoire avec une probabilité initiale u_0 :

temps	page 1	page 2	page 3	page 4	page 5	page 6	page 7	page 8	page 9	page 10	page 11	page 12	page 13	page 14
$t=0$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000
$t=1$	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$t=2$	0.000	0.000	0.000	0.000	0.000	0.000	0.333	0.333	0.333	0.000	0.000	0.000	0.000	0.000
$t=3$	0.167	0.000	0.000	0.000	0.000	0.333	0.000	0.333	0.000	0.167	0.000	0.000	0.000	0.000
$t=4$	0.000	0.033	0.033	0.033	0.033	0.400	0.111	0.111	0.111	0.000	0.033	0.033	0.033	0.033
$t=5$	0.122	0.017	0.017	0.017	0.017	0.111	0.133	0.244	0.133	0.122	0.017	0.017	0.017	0.017
$t=6$	0.100	0.033	0.033	0.033	0.033	0.293	0.037	0.170	0.037	0.100	0.033	0.033	0.033	0.033
$t=7$	0.084	0.036	0.036	0.036	0.036	0.210	0.098	0.135	0.098	0.084	0.036	0.036	0.036	0.036
$t=8$	0.122	0.035	0.035	0.035	0.035	0.168	0.070	0.168	0.070	0.122	0.035	0.035	0.035	0.035
$t=9$	0.105	0.042	0.042	0.042	0.042	0.217	0.056	0.126	0.056	0.105	0.042	0.042	0.042	0.042
...														
$t=28$	0.125	0.050	0.050	0.050	0.050	0.151	0.050	0.100	0.050	0.125	0.050	0.050	0.050	0.050
$t=29$	0.125	0.050	0.050	0.050	0.050	0.150	0.050	0.100	0.050	0.125	0.050	0.050	0.050	0.050
$t=30$	0.125	0.050	0.050	0.050	0.050	0.150	0.050	0.100	0.050	0.125	0.050	0.050	0.050	0.050

On observe un phénomène de diffusion, plausible après réflexion.

Convergence : Après 30 itérations, on est très proche (à 10^{-3} près) de la solution μ déjà exhibée. Il ne s'agit pas de l'équiprobabilité : certaines pages sont plus fréquentées que d'autres !

Raffinement du modèle

Notre modèle a encore un défaut, illustré par l'exemple suivant :

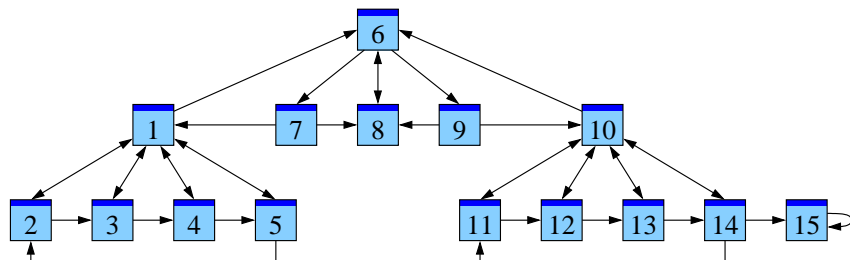


FIG.: Une variante du graphe initial

Raffinement du modèle

Notre modèle a encore un défaut, illustré par l'exemple suivant :

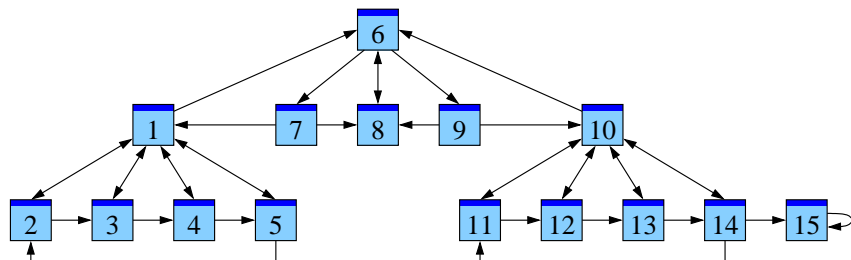


FIG.: Une variante du graphe initial

Google utilise un modèle raffiné dépendant d'un paramètre $c \in [0, 1]$:

Raffinement du modèle

Notre modèle a encore un défaut, illustré par l'exemple suivant :

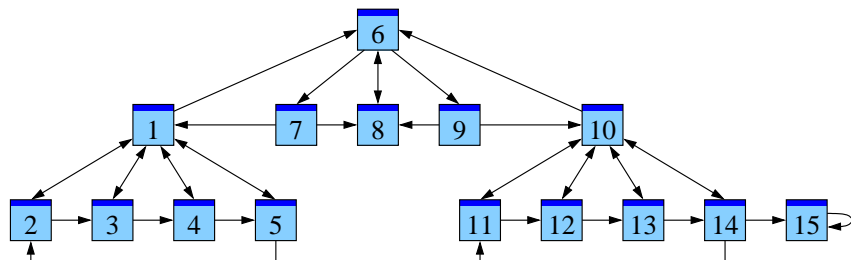


FIG.: Une variante du graphe initial

Google utilise un modèle raffiné dépendant d'un paramètre $c \in [0, 1]$:

- Avec probabilité c , le surfeur abandonne la page actuelle et recommence sur une des n pages du web, choisie au hasard.

Raffinement du modèle

Notre modèle a encore un défaut, illustré par l'exemple suivant :

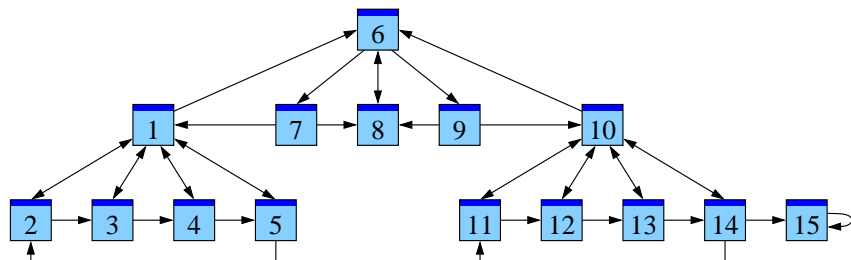


FIG.: Une variante du graphe initial

Google utilise un modèle raffiné dépendant d'un paramètre $c \in [0, 1]$:

- Avec probabilité c , le surfeur abandonne la page actuelle et recommence sur une des n pages du web, choisie au hasard.
- Avec probabilité $1 - c$, le surfeur suit un des liens de la page actuelle j , choisi de manière équiprobable, comme discuté avant.

Raffinement du modèle

Notre modèle a encore un défaut, illustré par l'exemple suivant :

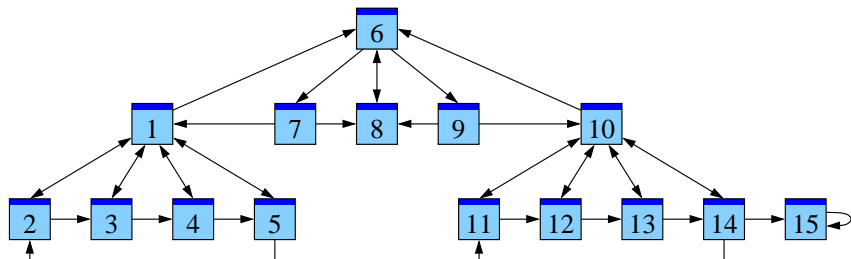


FIG.: Une variante du graphe initial

Google utilise un modèle raffiné dépendant d'un paramètre $c \in [0, 1]$:

- Avec probabilité c , le surfeur abandonne la page actuelle et recommence sur une des n pages du web, choisie au hasard.
- Avec probabilité $1 - c$, le surfeur suit un des liens de la page actuelle j , choisi de manière équiprobable, comme discuté avant.

Cette astuce évite de se faire piéger par une page sans issue.

Elle garantit d'arriver n'importe où dans le graphe, connexe ou non.

Le modèle PageRank utilisé par Google

Le modèle raffiné se formalise comme l'application affine

$$T: \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad x \mapsto c\varepsilon + (1 - c)Ax.$$

Le modèle PageRank utilisé par Google

Le modèle raffiné se formalise comme l'application affine

$$T: \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad x \mapsto c\varepsilon + (1 - c)Ax.$$

Ici A est la matrice stochastique définie ci-dessus.

Le modèle PageRank utilisé par Google

Le modèle raffiné se formalise comme l'application affine

$$T: \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad x \mapsto c\varepsilon + (1 - c)Ax.$$

Ici A est la matrice stochastique définie ci-dessus.

Le vecteur $\varepsilon = (\frac{1}{n}, \dots, \frac{1}{n})$ correspond à l'équiprobabilité.

La constante $c \in [0, 1]$ est un paramètre du modèle.

Le modèle PageRank utilisé par Google

Le modèle raffiné se formalise comme l'application affine

$$T: \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad x \mapsto c\varepsilon + (1 - c)Ax.$$

Ici A est la matrice stochastique définie ci-dessus.

Le vecteur $\varepsilon = (\frac{1}{n}, \dots, \frac{1}{n})$ correspond à l'équiprobabilité.

La constante $c \in [0, 1]$ est un paramètre du modèle.

Existence, unicité, et calcul de la solution

Sur \mathbb{R}^n on considère la norme $|x| = \sum_{k=1}^n |x_k|$.

Existence, unicité, et calcul de la solution

Sur \mathbb{R}^n on considère la norme $|x| = \sum_{k=1}^n |x_k|$.

Théorème

Soit $A \in \mathbb{R}^{n \times n}$ une matrice stochastique, soit $c \in]0, 1]$ une constante, et soit $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ l'application affine définie par

$$T(x) = c\varepsilon + (1 - c)Ax$$

Existence, unicité, et calcul de la solution

Sur \mathbb{R}^n on considère la norme $|x| = \sum_{k=1}^n |x_k|$.

Théorème

Soit $A \in \mathbb{R}^{n \times n}$ une matrice stochastique, soit $c \in]0, 1]$ une constante, et soit $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ l'application affine définie par

$$T(x) = c\varepsilon + (1 - c)Ax$$

Alors T est contractante de rapport $k = 1 - c < 1$.

Existence, unicité, et calcul de la solution

Sur \mathbb{R}^n on considère la norme $|x| = \sum_{k=1}^n |x_k|$.

Théorème

Soit $A \in \mathbb{R}^{n \times n}$ une matrice stochastique, soit $c \in]0, 1]$ une constante, et soit $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ l'application affine définie par

$$T(x) = c\varepsilon + (1 - c)Ax$$

Alors T est contractante de rapport $k = 1 - c < 1$.

Elle admet une unique mesure invariante $\mu = T(\mu)$ et pour tout vecteur initial u_0 la suite itérative $u_{m+1} = T(u_m)$ converge vers μ .

Existence, unicité, et calcul de la solution

Sur \mathbb{R}^n on considère la norme $|x| = \sum_{k=1}^n |x_k|$.

Théorème

Soit $A \in \mathbb{R}^{n \times n}$ une matrice stochastique, soit $c \in]0, 1]$ une constante, et soit $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ l'application affine définie par

$$T(x) = c\varepsilon + (1 - c)Ax$$

Alors T est contractante de rapport $k = 1 - c < 1$.

Elle admet une unique mesure invariante $\mu = T(\mu)$ et pour tout vecteur initial u_0 la suite itérative $u_{m+1} = T(u_m)$ converge vers μ .

Méthode de la puissance : Partant d'un vecteur stochastique initial, on itère l'application T . La proposition garantit la convergence.

Existence, unicité, et calcul de la solution

Sur \mathbb{R}^n on considère la norme $|x| = \sum_{k=1}^n |x_k|$.

Théorème

Soit $A \in \mathbb{R}^{n \times n}$ une matrice stochastique, soit $c \in]0, 1]$ une constante, et soit $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ l'application affine définie par

$$T(x) = c\varepsilon + (1 - c)Ax$$

Alors T est contractante de rapport $k = 1 - c < 1$.

Elle admet une unique mesure invariante $\mu = T(\mu)$ et pour tout vecteur initial u_0 la suite itérative $u_{m+1} = T(u_m)$ converge vers μ .

Méthode de la puissance : Partant d'un vecteur stochastique initial, on itère l'application T . La proposition garantit la convergence.

Vitesse. On a $|u_m - \mu| \leq k^m |u_0 - \mu|$, la convergence vers μ est donc au moins aussi rapide que celle de la suite géométrique k^m vers 0.

Résumé

- 1** Résolution de systèmes d'équations linéaires
 - Systèmes d'équations linéaires, l'algorithme de Gauss
 - Calcul matriciel : addition, multiplication, inversion, déterminant
 - Stabilité numérique, conditionnement d'une matrice
- 2** Réduction des endomorphismes
 - Espaces vectoriels et applications linéaires
 - Vecteurs propres, polynôme caractéristique
 - Polynôme minimal, méthodes de calcul
- 3** Méthodes approchées itératives
 - La méthode de la puissance
 - La méthode des itérations inverses
 - Matrices hermitiennes et symétriques
- 4** Comment fonctionne Google ?
 - Comment mesurer l'importance d'une page web ?
 - Le modèle PageRank : marche aléatoire sur le web
 - Existence, unicité, et calcul de la solution