

Problème discret

Systèmes linéaires

Difficult

Lacinpics

données

Algèbre linéaire Rappels

Référence

## MODÉLISATION ET INVERSION EN GÉOPHYSIQUE 4 - Inversion: Introduction

Bernard Giroux
(bernard.giroux@ete.inrs.ca)

Institut national de la recherche scientifique Centre Eau Terre Environnement

> Version 1.0.5 Hiver 2018



oblème discre

Systèmes linéaires

Difficulté

Exemple

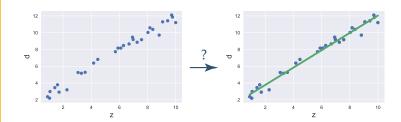
Nature aléatoire de données

Algèbre linéaire Rappels

Référence

### Données et modèles







# Aperçu

#### Données et modèles

oblème disc

Systèmes linéaire

Difficult

Exemple

données

Algèbre linéaire Rappels

Références

• Problème direct

Problème inverse

#### Données et modèles

Problème discret

Systèmes linéair

Difficult

Exemp

Nature aléatoire c données

Algèbre linéaire Rappels

Références

- Les données mesurées, notées *d*, constituent le point de départ de l'inversion.
- L'objectif est d'obtenir une information caractérisant l'objet étudié;
  - Cette information prends la forme de valeurs numériques : les paramètres du modèle, notés *m*.
- Les lois de la physique permettent de relier *m* et *d*;
  - Ces lois sont décrites par une fonction *G*, telle que

$$G(m) = d. (1)$$

 Les données peuvent être fonction du temps et/ou de l'espace, et sont généralement une série d'observations discrètes.



### Données et modèles

roblème dis

Systèmes linéain

Difficulté

Exemples

nnées

Rappels

D ( ( ) ....

- En pratique, les données mesurées contiennent une erreur expérimentale.
- On assume que les données sont la somme des mesures obtenues d'une expérience "parfaite", notées  $d_{\text{vrai}}$ , et d'un bruit  $\eta$ , i.e.

$$d = G(m_{\text{vrai}}) + \eta \tag{2}$$

$$=d_{\text{vrai}}+\eta,\tag{3}$$

où

- $d_{\text{vrai}}$  satisfait l'éq. (1) lorsque m est égal au modèle vrai  $m_{\text{vrai}}$ ;
- la fonction *G* représente exactement la réalité.
- La présence de  $\eta$ , même faible, peut faire en sorte que m retrouvé par inversion soit très différent de  $m_{\text{vrai}}$ .
- En général, il existe un infinité de modèles m différents de  $m_{\rm vrai}$  qui s'ajustent à  $d_{\rm vrai}$ .



Problème discret

Systèmes linéaires

Difficulte

Exemple

Nature aléatoire de données

Algèbre linéaire Rappels

Référence

# Problème discret



### Problème discret

Domices et modere.

#### Problème discret

Systèmes li

Difficult

#### Exemple

Vature aléatoire c données

Rappels

Références

• La plupart du temps, le modèle est décrit par un nombre fini, *M*, de paramètres, i.e.

$$\mathbf{m} = [m_0, m_1, m_2, \dots, m_{M-1}] \tag{4}$$

De façon similaire, on dispose d'un nombre fini, *N*, de données

$$\mathbf{d} = [d_0, d_1, d_2, \dots, d_{N-1}] \tag{5}$$

• On a alors affaire à un problème inverse discret de la forme

$$G(\mathbf{m}) = \mathbf{d}.\tag{6}$$

- Dans le cas contraire où le modèle et les données sont des fct continues, l'estimation de m à partir de d est un problème inverse continu;
  - On peut souvent approximer un problème continu par un problème discret.



### Problème discret

Données et modèles

#### Problème discret

Systèmes linéaire

Difficult

#### Exemples

données

Rappels

- Lorsque le nombre de paramètres *M* est faible, on parle d'estimation de paramètres;
- A contrario, lorsque M est élevé et qu'il est nécessaire d'appliquer des contraintes pour stabiliser la solution, on parle de problème inverse;
  - On verra plus loin que des contraintes sont nécessaires lorsque le système à résoudre est *mal conditionné*.



oblème disci

Systèmes linéaires

Difficulté

Exemples

Nature aléatoire de données

Algèbre linéaire Rappels

Référence

# Systèmes linéaires

# Systèmes linéaires

Données et modèle

roblème disc

#### Systèmes linéaires

Difficult

Exemple

données

Rappels

Ráfárancas

- Les systèmes linéaires sont un type de modèle mathématique trouvant plusieurs applications;
- Les systèmes linéaires obéissent au principe de superposition :

$$G(m_1 + m_2) = G(m_1) + G(m_2)$$
 (7)

et à la mise à l'échelle :

$$G(\alpha m) = \alpha G(m). \tag{8}$$

 Dans le cas des problèmes inverses discrets, le problème devient un système linéaire d'équations algébriques :

$$G(\mathbf{m}) = \mathbf{G}\mathbf{m} = \mathbf{d}.\tag{9}$$

où **G** est de taille  $N \times M$ .



## Systèmes linéaires

Données et modèles

oblème disci

#### Systèmes linéaires

Difficult

données

Rappels

- En géophysique, les modèles linéaires sont souvent utilisables;
- La raison principale est que l'objet d'étude varie peu par rapport à son état d'équilibre;
- Une relation linéaire permet de décrire adéquatement le phénomène;
- Par exemple en sismique, les contraintes générées par le passage des ondes sont très faibles p/r aux modules d'élasticité;
  - La relation contrainte/déformation est alors quasi linéaire.
- La gravimétrie et le magnétisme sont d'autres exemples où les champs sont faibles et où des modèles linéaires s'appliquent.



blème discr

Systèmes linéaires

#### Difficultés

Exemple

Nature aléatoire de données

Algèbre linéaire Rappels

Dáfáranca

## Difficultés



## Difficultés du problème inverse

Données et modèles

oblème disci

Systèmes linéaire

#### Difficultés

Exemp

Nature aléatoire c données

Rappels

- Il est crucial de demeurer critique face aux résultats de l'inversion;
- La raison principale est qu'il peut y avoir plusieurs modèles qui s'ajustent aussi bien aux données;
- Les éléments à l'origine de ce phénomène sont :
  - l'existence d'une solution;
  - 2 la non unicité de la solution;
  - 3 l'instabilité du système.



### Existence de la solution

Données et modèles

oblème disc

Systèmes linéaire

#### Difficultés

Exemples

données

Rappels

- Il est possible qu'aucun modèle ne s'ajuste *parfaitement* aux données;
- Les raisons sont :
  - le modèle physique est approximatif;
  - les données contiennent du bruit.
- Si l'ajustement n'est pas parfait, il est fort probable que le modèle estimé ne soit qu'une approximation du modèle réel.



### Non unicité de la solution

Données et modèles

Problème discret

Systèmes linéaire

#### Difficultés

------

données

Rappels

- Advenant que des solutions exactes existent, elles peuvent être non uniques, *même pour un nombre infini de données*;
- L'exemple classique est la réponse d'une sphère en gravimétrie, qui dépend de la masse de la sphère et non de la distribution de densité.
  - Deux sphères donneront exactement la même réponse si

$$\rho_1 \frac{4}{3} \pi r_1^3 = \rho_2 \frac{4}{3} \pi r_2^3$$

- La non unicité est une caractéristique des systèmes linéaires pour lesquelles les équations ne sont pas toutes linéairement indépendantes;
  - Le degré d'indépendance peut être évalué par l'analyse de la résolution du modèle.



### Instabilité

Données et modèles

roblème disc

Systèmes linéaire

#### Difficultés

Exemples

Nature aléatoire d données

Algèbre linéaire Rappels

Dáfárancac

- Une solution est instable lorsqu'un faible changement dans une mesure (e.g., un faible bruit  $\eta$ ) produit une variation importante du modèle estimé;
- De tels problèmes sont dits *mal conditionnés* dans le cas des problèmes discrets, ou *mal posés* dans le cas continu;
- Il est possible de stabiliser la solution en imposant des contraintes qui vont biaiser (d'une façon souhaitée) la solution;
  - on parle alors de régularisation.



oblème discr

Systèmes linéaires

Difficulté

#### Exemples

Nature aléatoire de données

Algèbre linéaire Rappels

Référence

# Exemples



### Exemple 1: Ajuster une droite

Données et modèles

Problème discre

Systèmes linéaire

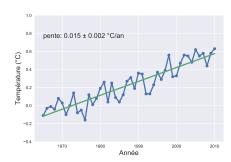
Difficultes

### Exemples

nnées

Algèbre linéaire Rappels

- On dispose d'un certain nombre (*N*) de mesures de température prises à des temps *t<sub>i</sub>* dans l'atmosphère.
  - Ces données contsituent le vecteur  $\mathbf{d} = [T_0, T_1, T_2, \dots, T_{N-1}]^T$ .
- On assume que la température obéit à un modèle linéaire en fonction du temps : T = a + bt;
  - L'ordonnée à l'origine a et la pente b sont les deux paramètres du modèle, i.e.  $\mathbf{m} = [a, b]^T$ .





## Exemple 1 : Ajuster une droite

Donnees et modeles

roblème dis

Systèmes linéai

Difficulte

### Exemples

données

Algèbre linéaire Rappels

Références

• Selon le modèle linéaire, la température doit satisfaire

$$T_0 = a + bt_0 \tag{10}$$

$$T_1 = a + bt_1 \tag{11}$$

$$T_{N-1} = a + bt_{N-1} (13)$$

Sous forme matricielle, on a

$$\begin{bmatrix}
T_0 \\
T_1 \\
\vdots \\
T_{N-1}
\end{bmatrix} = \begin{bmatrix}
1 & t_0 \\
1 & t_1 \\
\vdots & \vdots \\
1 & t_{N-1}
\end{bmatrix} \underbrace{\begin{bmatrix} a \\ b \end{bmatrix}}_{\mathbf{m}}$$
(14)



## Exemple 2: Profilage sismique vertical

Données et modèles

Duolal Sanco di como

Systèmes linéair

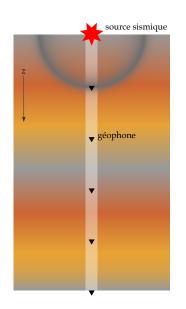
Difficultés

#### Exemples

données

Algèbre linéaire Rappels

- Avec le profilage sismique vertical, on cherche à déterminer la distribution verticale de la vitesse sismique V;
- Des géophones sont placés dans un forage et une source est actionnée à la surface;
- L'onde sismique est enregistrée aux géophones, ce qui permet de déterminer le temps de parcours t.





# **Exemple 2 : Profilage sismique vertical**

Données et modèles

roblème disci

Systèmes linéair

Difficulté

### Exemples

données

Rappels

Diffman

- Le problème est non linéaire lorsque défini en terme de vitesse;
- Le problème devient linéaire si exprimé en terme de *lenteur* (s), l'inverse de la vitesse, i.e. s = 1/V.
- Le temps de parcours à une profondeur z vaut

$$t(z) = \int_{0}^{z} s(l)dl \tag{15}$$

$$= \int_0^\infty s(l)H(z-l)dl \tag{16}$$

où H est la fonction de Heaviside, qui vaut 1 si  $z-l \geq 0$  et 0 si z-l < 0.



# Exemple 2: Profilage sismique vertical

Données et modèles

roblème discr

Systèmes linéair

Difficulte

### Exemples

onnées

Rappels

Diffman

- Le problème est résolu en discrétisant le milieu en couches
- Si le modèle compte M couches et le levé compte N géophones, l'intégrale devient, pour un  $i^e$  géophone à une position  $y_i$

$$t_{i} = \sum_{j=0}^{M-1} H(y_{i} - z_{j}) s_{j} \Delta z$$
 (17)

où  $M/N = \Delta y/\Delta z$  est un entier.

- Le vecteur des données est  $\mathbf{d} = [t_0, t_1, t_2, \dots, t_{N-1}]^T$ ;
- Les paramètres du modèle est sont regroupés dans le vecteur  $\mathbf{m} = [s_0, s_1, s_2, \dots, s_{M-1}]^T$ ;
- La matrice **G** sera alors de dimension  $N \times M$  et contiendra les termes  $H(y_i z_j)\Delta z$ .



# **Exemple 2 : Profilage sismique vertical**

Données et modèles

oblème discr

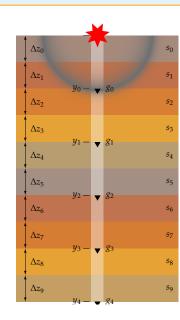
Systèmes linéaire

Difficult

#### Exemples

données

Algèbre linéai Rappels





# **Exemple 2 : Profilage sismique vertical**

Données et modèles

oblème discret

Systèmes linéair

Difficult

### Exemples

données

Rappels

Références

- Pour le cas illustré à la figure précédente, M/N = 2.
- Sous forme matricielle, pour l'ensemble des données de la figure, on obtient

• Chaque ligne i contient  $(i+1)\cdot M/N$  éléments égaux à  $\Delta z$  et  $M-(i+1)\cdot M/N$  zéros.



## Exemple 3: Tomographie

Données et modèles

roblème discre

Systèmes linéaire

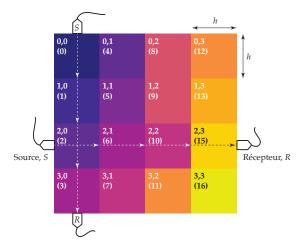
Difficultes

#### Exemples

Nature aléatoire de données

Algèbre linéair Rappels

- En tomographie, on cherche à évaluer la vitesse de propagation ou l'atténuation des ondes dans un milieu.
- Soit l'exemple d'un mur de briques de vitesses différentes :





# **Exemple 3: Tomographie**

Données et modèles

roblème discr

Systèmes linéair

Dimeune

### Exemples

données

Algèbre linéaire Rappels

- Deux séries de mesures sont faites, une première le long des colonnes et la seconde le long des lignes, pour un total de N=8 mesures.
- Le vecteur des données est  $\mathbf{d} = [t_0, t_1, t_2, \dots, t_7]^T$
- On suppose que chaque brique est de vitesse *V* uniforme;
- Le temps de parcours dans une brique j est proportionnel à la distance parcourue dans la brique, h, et vaut  $t_j = hs_j$ , où s est la lenteur.
  - Le modèle comporte M=16 paramètres, et est dans ce cas  $\mathbf{m} = [s_0, s_1, s_2, \dots, s_{15}]^T$

# Exemple 3: Tomographie

Données et modèle

oblème disc

Systèmes linéair

Difficulté

### Exemples

Nature aléatoire o données

Algèbre linéair Rappels

Référence

• On relie les données aux paramètres du modèle par

colonne 1: 
$$t_0 = hs_0 + hs_1 + hs_2 + hs_3$$
  
colonne 2:  $t_1 = hs_4 + hs_5 + hs_6 + hs_7$   
:  
ligne 3:  $t_6 = hs_2 + hs_6 + hs_{10} + hs_{14}$   
ligne 4:  $t_7 = hs_3 + hs_7 + hs_{11} + hs_{15}$ 

• Sous forme matricielle, nous avons

(19)



roblème discr

Systèmes linéaires

Difficulte

Exemple

Nature aléatoire des données

Algèbre linéaire Rappels

Référence

## Nature aléatoire des données



### Nature aléatoire des données

Données et modèles

Problème discre

Systèmes line

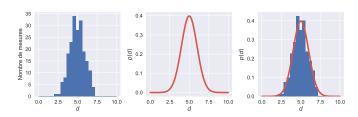
Difficultés

Exemple

Nature aléatoire des données

Algèbre linéaire Rappels

- Invariablement, une mesure expérimentale contient du bruit;
- Une observation répétée au même point donnera des mesures différentes;
  - On dit de la variable observée que c'est une variable aléatoire, et chaque mesure est une réalisation de cette variable aléatoire;
- Une variable aléatoire possède des propriétés précises qui dictent la *distribution* des valeurs observées;
  - Les réalisations permettent *d'estimer* ces propriétés.





### Nature aléatoire des données

Données et modèles

Problème discre

Systèmes linéair

Difficulté

Exemples

Nature aléatoire des données

Rappels

- Les propriétés d'une variables aléatoire d sont spécifiées par sa fonction de densité de probabilité (f.d.p), notée p(d);
  - Cette fonction donne la probabilité qu'une réalisation aura une valeur au voisinage de *d*.
- La fonction p(d) est souvent compliquée et ne peut pas être évaluée directement.
- On résume plutôt ses caractéristiques principales avec quelques grandeurs particulières, par exemple :
  - la moyenne, notée  $\langle d \rangle$ ;
  - la variance, notée  $\sigma^2$ .



Données et modèles

Problème discre

Systèmes linéair

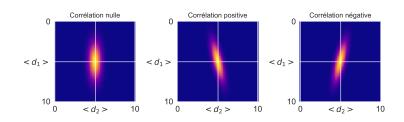
Difficultés

Exemple

Nature aléatoire des données

Algèbre linéaire Rappels

- Les levés géophysiques sont réalisés en prenant des mesures en plusieurs points;
- Il peut arriver que les mesures soient corrélées;
  - Des valeurs élevées en un point surviennent de façon consistante avec d'autre valeurs élevées (ou faibles) en un autre point.





Données et modèles

roblème discr

Systèmes linéair

Difficulté:

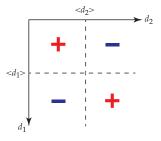
Exempl

Nature aléatoire des données

Algèbre linéaire Rappels

Référence

• Le degré de corrélation de deux variables  $d_1$  et  $d_2$  peut être quantifié en séparant la f.d.p. conjointe en 4 quadrants;



• Si les variables ne sont pas corrélées, la somme des valeurs de la f.d.p. conjointe de chaque quadrant sera nulle.



Données et modèles

roblème discr

Systèmes linéair

Difficulté

Exemples

Nature aléatoire des données

Rappels

Références

- Si la fonction permettant de construire les quadrants est  $[d_1 \langle d_1 \rangle][d_2 \langle d_2 \rangle]$ , la mesure du degré de corrélation est appelée covariance;
- Pour un ensemble de N données ayant chacune K réalisations regroupées dans une matrice  $\mathbf{D}$  de taille  $K \times N$ , la covariance expérimentale vaut

$$[\operatorname{cov} \mathbf{d}]_{ij}^{\operatorname{est}} = \frac{1}{K} \sum_{k=0}^{K-1} \left( D_{ki} - \langle D_i \rangle^{\operatorname{est}} \right) \left( D_{kj} - \langle D_j \rangle^{\operatorname{est}} \right). \tag{20}$$

où  $\langle D_i \rangle^{\text{est}}$  est la moyenne expérimentale de la  $i^e$  donnée.



Données et modèles

Problème discret

Systèmes linéaire

Difficultés

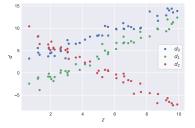
Exemples

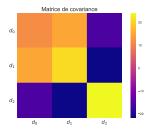
Nature aléatoire des données

Algèbre linéaire Rappels

Référence

• Exemple de matrice de covariance pour trois variables aléatoires :







Algèbre linéaire -Rappels

Algèbre linéaire – Rappels



# Algèbre linéaire – Rappels

Algèbre linéaire -Rappels

• Un système linéaire peut être résolu en utilisant l'élimination de Gauss-Jordan:

Soit le système

 $x_1 + 2x_2 + 3x_3 = 14$ (21)

 $x_1 + 2x_2 + 2x_3 = 11$ (22)

 $x_1 + 3x_2 + 4x_3 = 19$ 

En soustrayant la  $1^{re}$  équation aux deux autres :

$$x_1 + 2x_2 + 3x_3 = 14$$

$$-x_3 = -3 (25)$$
$$x_2 + x_3 = 5 (26)$$

• On interchange les 
$$2^e$$
 et  $3^e$  équations

$$x_1 + 2x_2 + 3x$$

$$x_1 + 2x_2 + 3x_3 = 14 \tag{27}$$

$$\begin{aligned}
 x_2 + x_3 &= 5 \\
 -x_3 &= -3 
 \end{aligned} 
 \tag{28}$$

(23)

(24)



# Algèbre linéaire – Rappels

Données et modèles

oblème disc

Systèmes linéair

Difficulté

Exemples

données

Algèbre linéaire – Rappels

• On élimine  $x_2$  de la  $1^{re}$  équation en soustrayant 2 fois la  $2^e$  (on change le signe de la  $3^e$  au passage...)

$$x_1 + x_3 = 4 (30)$$

$$x_2 + x_3 = 5 (31)$$

$$x_3 = 3 \tag{32}$$

• On élimine finallement  $x_3$  des deux premières équations pour arriver à la solution :

$$x_1 = 1 \tag{33}$$

$$x_2 = 2 \tag{34}$$

$$x_3 = 3 \tag{35}$$



# Algèbre linéaire – Rappels

Donnees et modeles

oblème disc

Systèmes linéair

Dimcuite

LACIN

données

Algèbre linéaire – Rappels

Définitions

Référeno

• On peut procéder avec une *matrice augmentée*, définie pour l'exemple précédent par

$$\begin{bmatrix}
1 & 2 & 3 & 14 \\
1 & 2 & 2 & 11 \\
1 & 3 & 4 & 19
\end{bmatrix}$$
(36)

• En effectuant les opérations de l'élimination de Gauss-Jordan sur les lignes de cette matrice, on arrive à

$$\begin{bmatrix}
1 & 0 & 0 & | & 1 \\
0 & 1 & 0 & | & 2 \\
0 & 0 & 1 & | & 3
\end{bmatrix}$$
(37)

et les termes de la colonne de droite sont la solution.



Données et modèles

oblème discr

Systèmes linéair

Difficult

Exempl

données

Rappels Définitions

\_\_\_\_\_

- Une matrice est dite échelonnée réduite en lignes (ERL) si :
  - le premier élément non nul sur chaque ligne est un 1. Cet élément est appelé un *élément pivot*, la colonne où apparaît cet élément est une *colonne pivot*;
  - À part l'élément pivot, tout les autres éléments d'une colonne sont des 0;
  - Chaque élément pivot est à droite de l'élément pivot des lignes précédentes;
  - Des lignes ne comportant que des 0 sont situées au bas de la matrice.
  - La résolution d'un système linéaire sous une forme de matrice augmentée consiste à la transformer en la forme ERL.

Données et modèles

oblème disc

Systèmes linéaire

Dimcuite

Lacinpics

lonnées

Rappels
Définitions

Définitions

• Des vecteurs  $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$  sont linéairement indépendants si le système

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_n\mathbf{v}_n = 0 \tag{38}$$

n'admet que la solution c = 0.

• Si les colonnes d'une matrice **A** sont linéairement indépendantes, alors l'inverse de **A** existe.



Données et modèles

roblème discr

Systèmes linéaire

Difficult

Exempi

Nature aléatoire d lonnées

Rappels

Définitions

- On note  $R^n$  un espace à n dimensions.
- Un sous-espace W de  $\mathbb{R}^n$  est un sous-ensemble de  $\mathbb{R}^n$  qui satisfait :
  - si x et y sont des vecteurs dans W, alors x + y est aussi un vecteur dans W;
  - si *x* est un vecteur dans *W* et *s* un scalaire réel, alors *sx* est aussi un vecteur dans *W*;
  - **0** étant un vecteur dans *W*, le sous-espace est non-trivial s'il contient d'auters vecteurs que **0**.



Données et modèles

roblème discr

Systèmes linéair

Difficult

Exemp

Nature aléatoire données

Algèbre linéaire Rappels

Définitions

- Soit **A** une matrice m par n. Le noyau de **A**, noté  $N(\mathbf{A})$ , est l'ensemble des vecteurs  $\mathbf{x}$  tels que  $\mathbf{A}\mathbf{x} = \mathbf{0}$ .
- On peut montrer que  $N(\mathbf{A})$  est un sous-espace de  $\mathbb{R}^n$ .
- La base du sous-espace W est un ensemble de vecteurs  $\mathbf{v}_1$ , ...,  $\mathbf{v}_p$  tels que :
  - n'importe quel vecteur dans W est une combinaison linéaire des vecteurs de base;
  - les vecteurs de base sont linéairement indépendants.
- Une base standard de  $R^n$  est un ensemble de vecteurs  $\mathbf{e}_1, \ldots, \mathbf{e}_n$  tels que tout les éléments de  $\mathbf{e}_i$  sont des zéros sauf le  $i^e$  élément qui vaut 1.



Données et modèles

Problème discret

Systèmes linéair

Difficult

Exemp

Vature aléatoire d Jonnées

Rappels
Définitions

- Soit W un sous-espace de  $\mathbb{R}^n$  avec la base  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p$ . Toutes les bases de W ont p vecteurs et p est la dimension de W.
- Soit A une matrice m par n. L'espace colonne de A, noté R(A), est l'ensemble des vecteurs b tels que Ax = b possède au moins une solution.
  - Autrement dit, l'espace colonne est l'ensemble des vecteurs b qui peuvent s'écrire comme une combinaison linéaire des colonnes de A.
- On peut montrer que R(A) correspond aux colonnes pivot de A et que N(A) correspond aux colonnes non pivot. Il découle que

$$\dim N(\mathbf{A}) + \dim R(\mathbf{A}) = n \tag{39}$$



blème discret

Systèmes linéaires

Difficulte

Exemple

Nature aléatoire de données

Algèbre linéaire Rappels

Références



### Références

Données et modèles

oblème disc

Systèmes linéair

Dimcuit

Naturo aléatoiro

Algèbre linéaire Rappels

- Aster, R. C., Borchers, B., and Thurber, C. H. (2013). *Parameter Estimation and Inverse Problems*. Academic Press, 2nd edition
- Menke, W. (2012). *Geophysical Data Analysis : Discrete Inverse Theory*. Academic Press, 3<sup>rd</sup> edition