Modélisation et inversion en géophysique
4 - Inversion: Introduction

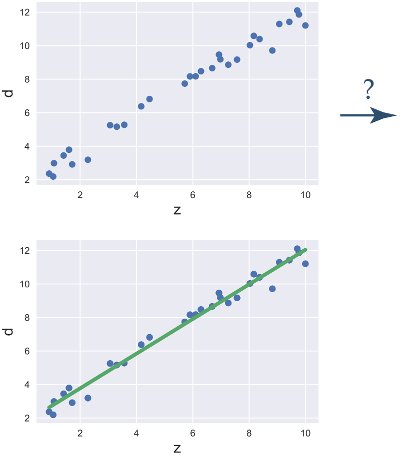
Bernard Giroux
(bernard.giroux@ete.inrs.ca)

Institut national de la recherche scientifique
Centre Eau Terre Environnement

Version 1.1.1
Hiver 2019

Données et modèles

Aperçu



Aperçu

- Problème direct

$\boxed{\text{estimé des paramètres du modèle}} \rightarrow \boxed{\text{modèle quantitatif}} \rightarrow \boxed{\text{prédiction des données}}$

- Problème inverse

$\boxed{\text{données observées}} \rightarrow \boxed{\text{modèle quantitatif}} \rightarrow \boxed{\text{estimé des paramètres du modèle}}$

Données et modèles

- 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

$$G(m) = d. \tag{1}$$

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

Données et modèles

- 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_{vrai}, et d'un bruit η, i.e.

$$
\begin{aligned}
d &= G(m_{\text{vrai}}) + \eta && (2)\\
&= d_{\text{vrai}} + \eta, && (3)
\end{aligned}
$$

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

Problème discret

Problème discret

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

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

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

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

- 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.

- 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*.

Systèmes linéaires

Systèmes linéaires

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

$$G(m_1 + m_2) = G(m_1) + G(m_2) \tag{7}$$

  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 $\mathbf{G}$ est de taille $N \times M$.

Systèmes linéaires

- 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.

Difficultés

Difficultés du problème inverse

- 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 :
  1. l'existence d'une solution ;
  2. la non unicité de la solution ;
  3. l'instabilité du système.

Existence de la solution

- 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

- 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*.

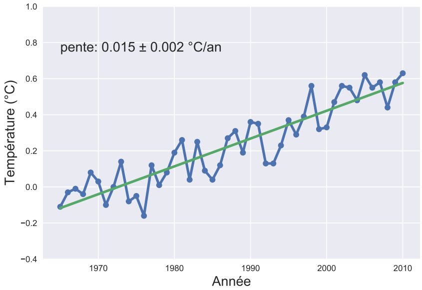
Instabilité

- Une solution est instable lorsqu'un faible changement dans une mesure (e.g., un faible bruit η) 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*.

Exemples

Exemple 1 : Ajuster une droite

- On dispose d'un certain nombre (N) de mesures de température prises à des temps t_i dans l'atmosphère.
  - Ces données contsituent le vecteur $\mathbf{d} = [T_0, T_1, T_2, \ldots, T_{N-1}]^T$.
- 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

- Selon le modèle linéaire, la température doit satisfaire

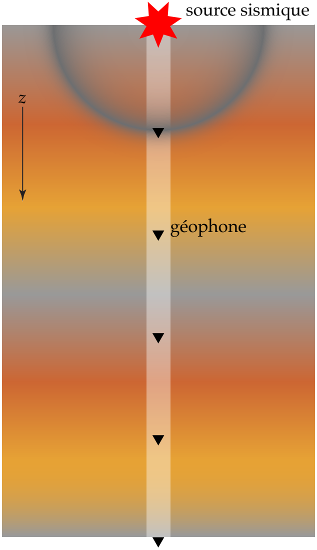
$$T_0 = a + bt_0 \tag{10}$$
$$T_1 = a + bt_1 \tag{11}$$
$$\vdots \tag{12}$$
$$T_{N-1} = a + bt_{N-1} \tag{13}$$

- Sous forme matricielle, on a

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

Exemple 2 : Profilage sismique vertical

- 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

- 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

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

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

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

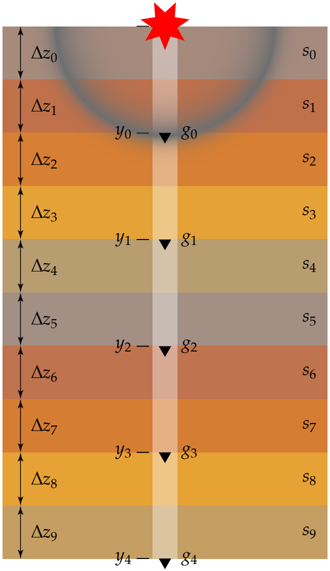
- Le problème est résolu en discrétisant le milieu en couches
- Si le modèle compte M couches et le levé compte N géophones, l'intégrale devient, pour un i^e géophone à une position y_i

$$t_i = \sum_{j=0}^{M-1} H(y_i - z_j) s_j \Delta z \tag{17}$$

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

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

Exemple 2 : Profilage sismique vertical



Exemple 2 : Profilage sismique vertical

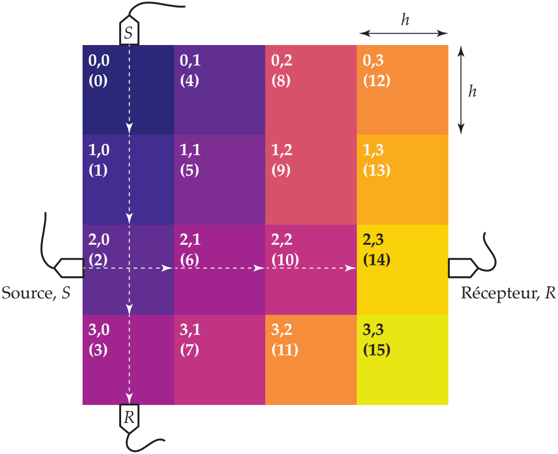
- Pour le cas illustré à la figure précédente, $M/N = 2$.
- Sous forme matricielle, pour l'ensemble des données de la figure, on obtient

$$\underbrace{\begin{bmatrix} t_0 \\ t_1 \\ t_2 \\ t_3 \\ t_4 \end{bmatrix}}_{\mathbf{d}} = \underbrace{\Delta z \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}}_{\mathbf{G}} \underbrace{\begin{bmatrix} s_0 \\ s_1 \\ s_2 \\ \vdots \\ s_9 \end{bmatrix}}_{\mathbf{m}} \tag{18}$$

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

Exemple 3 : Tomographie

- 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

- 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 $\mathbf{d} = [t_0, t_1, t_2, \ldots, t_7]^T$
- 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 $\mathbf{m} = [s_0, s_1, s_2, \ldots, s_{15}]^T$

Exemple 3 : Tomographie

- On relie les données aux paramètres du modèle par

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

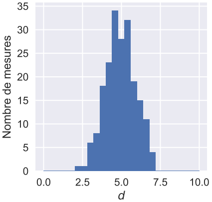
- Sous forme matricielle, nous avons

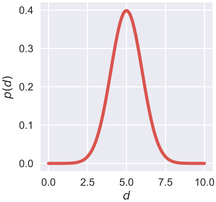
$$
\begin{bmatrix} t_0 \\ t_1 \\ \vdots \\ t_6 \\ t_7 \end{bmatrix} = h \begin{bmatrix} 1&1&1&1&0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&1&1&1&1&0&0&0&0&0&0&0&0 \\ \vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots \\ 0&0&1&0&0&0&1&0&0&0&1&0&0&0&1&0 \\ 0&0&0&1&0&0&0&1&0&0&0&1&0&0&0&1 \end{bmatrix} \begin{bmatrix} s_0 \\ s_1 \\ s_2 \\ \vdots \\ s_{15} \end{bmatrix} \tag{19}
$$

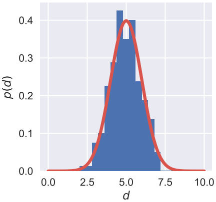
Nature aléatoire des données

Nature aléatoire des données

- 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.





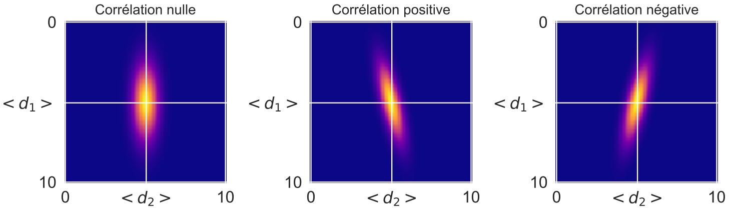


Nature aléatoire des données

- 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 σ^2.

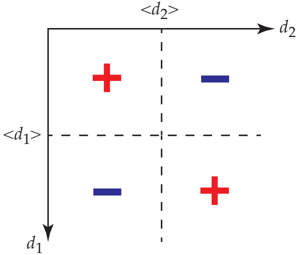
Données corrélées

- 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 corrélées

- 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 corrélées

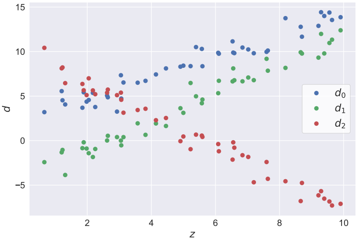
- Si la fonction permettant de construire les quadrants est $[d_1 - \langle d_1 \rangle][d_2 - \langle d_2 \rangle]$, la mesure du degré de corrélation est appelée covariance ;
- Pour un ensemble de N données ayant chacune K réalisations regroupées dans une matrice $\mathbf{D}$ de taille $K \times N$,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

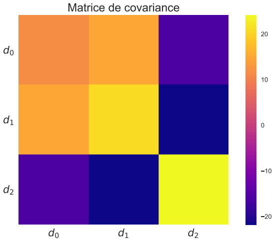
$$[\mathrm{cov}\,\mathbf{d}]_{ij}^{\mathrm{est}} = \frac{1}{K} \sum_{k=0}^{K-1} \left(D_{ki} - \langle D_i \rangle^{\mathrm{est}} \right) \left(D_{kj} - \langle D_j \rangle^{\mathrm{est}} \right). \qquad (20)$$

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

Données corrélées

- Exemple de matrice de covariance pour trois variables aléatoires :





Algèbre linéaire – Rappels

Algèbre linéaire – Rappels

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

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 14 \\ x_1 + 2x_2 + 2x_3 &= 11 \\ x_1 + 3x_2 + 4x_3 &= 19 \end{aligned}$$

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

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 14 \\ -x_3 &= -3 \\ x_2 + x_3 &= 5 \end{aligned}$$

  - On interchange les 2^e et 3^e équations

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 14 \\ x_2 + x_3 &= 5 \\ -x_3 &= -3 \end{aligned}$$

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...)

$$\begin{aligned} x_1 + x_3 &= 4 \\ x_2 + x_3 &= 5 \\ x_3 &= 3 \end{aligned}$$

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

$$\begin{aligned} x_1 &= 1 \\ x_2 &= 2 \\ x_3 &= 3 \end{aligned}$$

Algèbre linéaire – Rappels

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

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 14 \\ 1 & 2 & 2 & 11 \\ 1 & 3 & 4 & 19 \end{array}\right]$$

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

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

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

Algèbre linéaire – Définitions

- Une matrice est dite *échelonnée réduite en lignes* (ERL) [1] 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.

1. *reduced row echelon form* (RREF) en anglais

Algèbre linéaire – Définitions

- Soit le système à deux équations

$$\begin{aligned} x_1 + x_2 + x_3 &= 0 \\ x_1 + 2x_2 + 2x_3 &= 0 \end{aligned}$$

- La forme ERL du système est

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right],$$

ce qui revient à

$$\begin{aligned} x_1 &= 0 \\ x_2 + x_3 &= 0 \end{aligned}$$

- Clairement, x_1 doit être 0 ;
- Par contre, x_2 et x_3 ne sont pas déterminés ; on peut traiter x_3 comme une variable libre et alors $x_2 = -x_3$.

Algèbre linéaire – Définitions

- Une matrice identité $\mathbf{I}$ est un matrice comportant des 1 sur sa diagonale, les autres éléments étant des 0.
- Soit $\mathbf{A}$ une matrice n par m. Si une matrice $\mathbf{B}$ existe telle que

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I}, \tag{21}$$

  alors $\mathbf{B}$ est l'inverse de $\mathbf{A}$, et l'on note $\mathbf{B} = \mathbf{A}^{-1}$.
- Des vecteurs $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ sont linéairement indépendants si le système

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n = 0$$

  n'admet que la solution $\mathbf{c} = \mathbf{0}$.
- Si les colonnes d'une matrice $\mathbf{A}$ sont linéairement indépendantes, alors l'inverse de $\mathbf{A}$ existe.

Algèbre linéaire – Définitions

Exemple

- Soit la matrice

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

- On peut déterminer si les colonnes sont indépendantes à partir de la forme ERL de $\mathbf{A}\mathbf{c} = \mathbf{0}$, qui est

$$\left[\begin{array}{ccr|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

- Les solutions sont

$$\begin{aligned} c_1 &= c_3 \\ c_2 &= -2c_3 \\ 0c_3 &= 0 \end{aligned} \qquad \text{et donc} \quad \mathbf{c} = c_3 \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

Algèbre linéaire – Définitions

Exemple (suite)

- En posant $c_3 = 1$, on obtient la solution non nulle

$$\mathbf{c} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

- Puisque $\mathbf{c}$ est différent de $\mathbf{0}$, les colonnes de $\mathbf{A}$ sont linéairement dépendentes et $\mathbf{A}$ n'a pas d'inverse.

Algèbre linéaire – Définitions

- On note R^n un espace à n dimensions.
- Un sous-espace W de R^n est un sous-ensemble de R^n qui satisfait :
  - si $\mathbf{x}$ et $\mathbf{y}$ sont des vecteurs dans W, alors $\mathbf{x} + \mathbf{y}$ est aussi un vecteur dans W ;
  - si $\mathbf{x}$ est un vecteur dans W et s un scalaire réel, alors $s\mathbf{x}$ est aussi un vecteur dans W ;
  - $\mathbf{0}$ étant un vecteur dans W, le sous-espace est non-trivial s'il contient d'autres vecteurs que $\mathbf{0}$.
- Par exemple, dans R^3, le plan

$$x_1 + x_2 + x_3 = 0$$

est un sous-ensemble de R^n car il satisfait aux critères précédents.

Algèbre linéaire – Définitions

- Soit $\mathbf{A}$ une matrice m par n. Le *noyau* (*null space*) de $\mathbf{A}$, noté $N(\mathbf{A})$, est l'ensemble des vecteurs $\mathbf{x}$ tels que $\mathbf{Ax} = \mathbf{0}$.
- On peut montrer que $N(\mathbf{A})$ est un sous-espace de R^n.
- En inversion, les propriétés du noyau sont critiques car un noyau non-trivial entraîne la non-unicité de la solution.
- Le noyau d'une matrice $\mathbf{A}$ est déterminé en solutionnant $\mathbf{Ax} = \mathbf{0}$.

Algèbre linéaire – Définitions

- Posons que $\mathbf{A}$ vaut

$$\mathbf{A} = \begin{bmatrix} 3 & 1 & 9 & 4 \\ 2 & 1 & 7 & 3 \\ 5 & 2 & 16 & 7 \end{bmatrix}$$

- On trouve le noyau de $\mathbf{A}$ à partir de la forme ERL de la matrice augmentée :

$$\left[\begin{array}{cccc|c} 1 & 0 & 2 & 1 & 0 \\ 0 & 1 & 3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

- Les solutions sont

$$\mathbf{x} = x_3 \begin{bmatrix} -2 \\ -3 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ -1 \\ 0 \\ 1 \end{bmatrix}$$

Algèbre linéaire – Définitions

- N'importe quel vecteur dans $N(\mathbf{A})$ est une combinaison linéaire de ces vecteurs.
- Supposons que $\mathbf{b} = [22\ 17\ 39]^T$.
- Une solution possible au système $\mathbf{Ax} = \mathbf{b}$ est $\mathbf{p} = [1\ 2\ 1\ 2]^T$ et n'importe quelle combinaison

$$\mathbf{p} + x_3 \begin{bmatrix} -2 \\ -3 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ -1 \\ 0 \\ 1 \end{bmatrix}$$

sera aussi une solution.

- De fait, si $\mathbf{x}$ est dans $N(\mathbf{A})$, alors

$$\begin{aligned} \mathbf{A}(\mathbf{x}+\mathbf{p}) &= \mathbf{Ax} + \mathbf{Ap} \\ \mathbf{A}(\mathbf{x}+\mathbf{p}) &= \mathbf{0} + \mathbf{b} \\ \mathbf{A}(\mathbf{x}+\mathbf{p}) &= \mathbf{b} \end{aligned}$$

Algèbre linéaire – Définitions

- 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.

Algèbre linéaire – Définitions

- Soit W un sous-espace de R^n avec la base $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$. Toutes les bases de W ont p vecteurs et p est la *dimension* de W, notée $dim\ W$.
- Soit $\mathbf{A}$ une matrice m par n. L'*espace colonne* (*range*) de $\mathbf{A}$, noté $R(\mathbf{A})$, est l'ensemble des vecteurs $\mathbf{b}$ tels que $\mathbf{Ax} = \mathbf{b}$ possède au moins une solution.
  - Autrement dit, l'espace colonne est l'ensemble des vecteurs $\mathbf{b}$ qui peuvent s'écrire comme une combinaison linéaire des colonnes de $\mathbf{A}$.
  - L'espace colonne est important en inversion car $R(\mathbf{G})$ est constitué de tout les vecteurs $\mathbf{d}$ pour lesquels il y a un modèle $\mathbf{m}$ tel que $\mathbf{Gm} = \mathbf{d}$.
- On peut montrer que $R(\mathbf{A})$ correspond aux colonnes pivot de $\mathbf{A}$ et que $N(\mathbf{A})$ correspond aux colonnes non pivot. Il découle que
$$dim\ N(\mathbf{A}) + dim\ R(\mathbf{A}) = n$$
- Le rang d'une matrice $\mathbf{A}$ est la dimension de $R(\mathbf{A})$.

Références

Références

- 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, 3rd edition