INRS
UNIVERSITÉ DE RECHERCHE

Modélisation et inversion en géophysique
4 - Inversion: Introduction

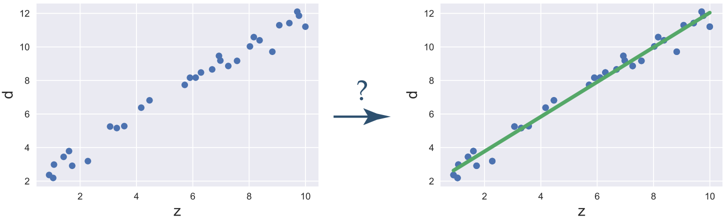
Bernard Giroux
(bernard.giroux@ete.inrs.ca)

Institut national de la recherche scientifique
Centre Eau Terre Environnement

Version 1.0.2
Hiver 2018

Données et modèles

Aperçu



Aperçu

- Problème direct

estimé des paramètres du modèle $\rightarrow$ modèle quantitatif $\rightarrow$ prédiction des données

- Problème inverse

données observées $\rightarrow$ modèle quantitatif $\rightarrow$ 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.

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*.

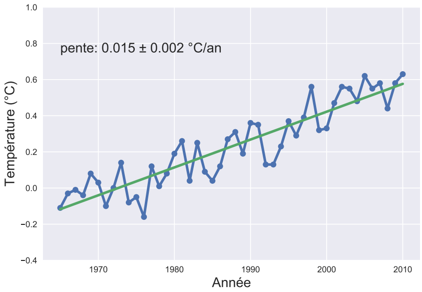
Instabilité

- 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

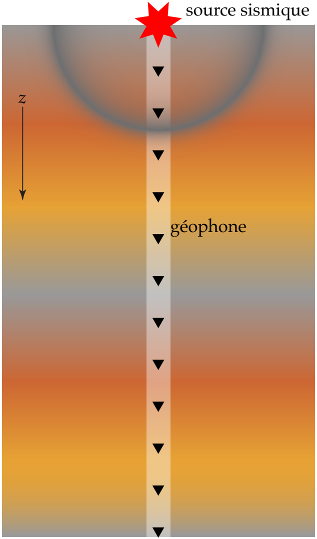
$$T_0 = a + bt_0 \tag{10}$$
$$T_1 = a + bt_1 \tag{11}$$
$$\vdots \tag{12}$$
$$T_{N-1} = a + bt_{N-1} \tag{13}$$

- 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

$$t(z) = \int_0^z s(l)dl \tag{15}$$

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

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

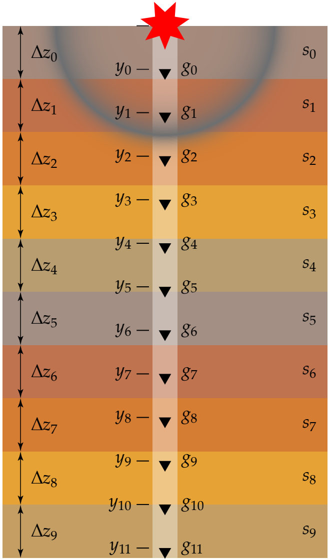
Exemple 2 : Profilage sismique vertical

- 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 $N/M = \Delta y / \Delta z$ 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



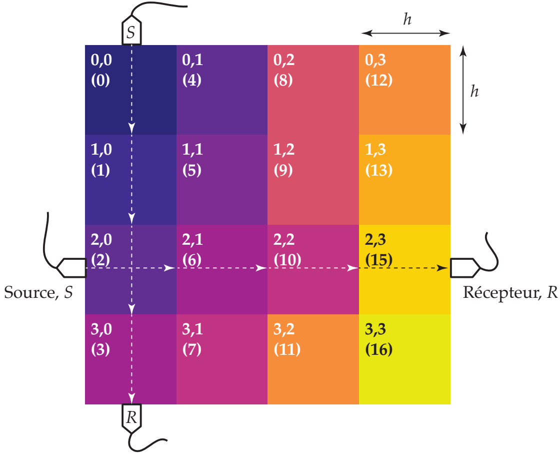
Exemple 2 : Profilage sismique vertical

- Sous forme matricielle, pour l'ensemble des données de la figure, on obtient

$$\underbrace{\begin{bmatrix} t_0 \\ t_1 \\ \vdots \\ t_{10} \\ t_{11} \end{bmatrix}}_{\mathbf{d}} = \underbrace{\Delta z \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}}_{\mathbf{G}} \underbrace{\begin{bmatrix} s_0 \\ s_1 \\ s_2 \\ \vdots \\ s_9 \end{bmatrix}}_{\mathbf{m}} \tag{18}$$

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

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

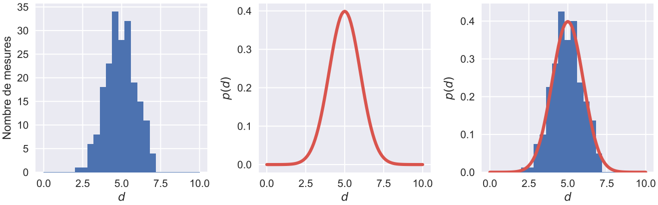
- Sous forme matricielle, nous avons

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

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.

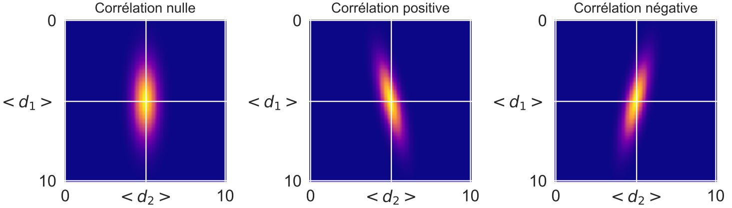


Nature aléatoire des données

- 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.

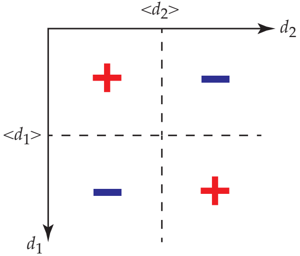
Données corrélées

- 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.

Références

INRS
UNIVERSITÉ DE RECHERCHE

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