OCEA0014-1 MÉTHODES MATHÉMATIQUES D'ANALYSE ET DE MODÉLISATION APPLIQUÉES À L'ENVIRONNEMENT

JANVIER 2015

Question I

La fonction d'Ivlev

$$f(x) = \mu_0 \left[1 - \mathrm{e}^{-x/a}\right]$$

où μ_0 et a sont des constantes strictement positives est fréquemment utilisée pour modéliser la réponse des systèmes biologiques.

i. *Déterminez une approximation linéaire de f valable au voisinage de $x = 0$.*

Une approximation linéaire au voisinage d'un point x_0 peut être générée en exploitant

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + o(x - x_0), \qquad (x \to x_0)$$

Appliquons ce résultat à la fonction d'Ivlev avec $x_0 = 0$. On calcule aisément $f(0) = 0$ et

$$f'(x) = \frac{\mu_0}{a}\,\mathrm{e}^{-x/a}, \qquad f'(0) = \frac{\mu_0}{a}$$

Dès lors,

$$f(x) = \frac{\mu_0 x}{a} + O(x), \qquad (x \to 0)$$

ii. *Déterminez le comportement asymptotique de f pour $x \to +\infty$.*

On calcule

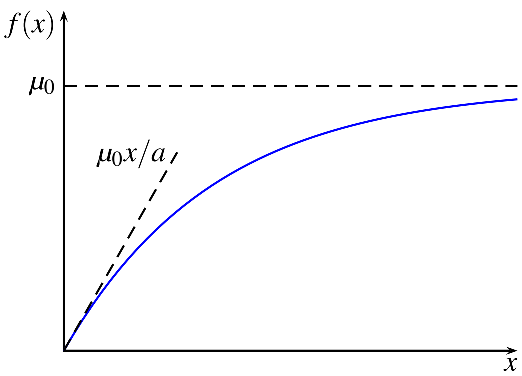
$$\lim_{x \to +\infty} \mu_0 \left[1 - \mathrm{e}^{-x/a}\right] = \mu_0$$

de sorte que

$$f(x) \sim \mu_0, \qquad (x \to +\infty)$$

iii. *Esquissez la fonction d'Ivlev pour les valeurs positives de x.*

Les considérations précédentes montrent que $f'(x) > 0$ quel que soit x. La fonction est donc strictement croissante et, compte tenu des comportements asymptotiques identifiés en i et ii, son graphe peut être esquissé de la façon suivante :



Question II

Dans leur article 'Mathematical modelling of seasonal migration with applications to climate change' à paraître prochainement dans Ecological Modelling, *Donohue et Piiroinen considèrent un modèle proche du modèle suivant décrivant la dynamique d'une population d'oiseaux migrateurs :*

$$\begin{cases} \dfrac{dM}{dt} = j - m_1 M \\ \dfrac{dB}{dt} = rM\left(1 - \dfrac{B}{\kappa}\right) - m_2 B \end{cases}$$

Les variables M et B désignent respectivement la biomasse des adultes et des juvéniles (exprimées en unités de masse). Le paramètre j désigne le flux d'adultes migrateurs vers la région considérée. Les paramètres j, m_1, r, κ et m_2 sont des constantes strictement positives.

i. *Donnez une interprétation possible des différents termes des équations de ce modèle.*

   Le terme j décrit l'arrivée d'adultes migrateurs à taux constant. Le terme $-m_1M$ en décrit la mortalité (ou la migration vers d'autres régions).

   La population des juvéniles croit de façon proportionnelle à la population des adultes dont ils représentent la progéniture, ce qui se traduit par le facteur rM. Cependant, l'espace pour la nidification des adultes est limité, de sorte que la croissance est limitée par un facteur $(1-B/\kappa)$ semblable au terme de croissance logistique. La population des juvéniles décroît également en fonction de leur mortalité ($-m_2B$).

ii. *Déterminez les dimensions des différents paramètres j, m_1, r, κ et m_2 et montrez que les équations du modèle peuvent être écrites sous la forme adimensionnelle*

$$\begin{cases} \dfrac{dx}{d\tau} = \alpha(1-x) \\ \dfrac{dy}{d\tau} = \beta x(1-y) - \gamma y \end{cases} \tag{†}$$

   en posant

$$x = \frac{m_1 M}{j}, \qquad y = \frac{B}{\kappa}, \qquad \tau = rt$$

   et en introduisant les nombres sans dimension α, β et γ. Déterminez l'expression des paramètres α, β et γ en fonction de j, m_1, r, κ et m_2.

   D'après l'énoncé, on sait que les populations des adultes et des juvéniles sont exprimées en unités de masse. On a donc

$$[M] = [B] = M$$

   Par le principe d'homogénéité dimensionnelle des équations, on a

$$[j] = [m_1 M] = \left[\frac{dM}{dt}\right] = MT^{-1} \qquad \text{et} \qquad [m_1] = T^{-1}$$

   De même,

$$[\kappa] = [B] = M$$

   et

$$[rM] = [m_2 M] = \left[\frac{dB}{dt}\right] = MT^{-1} \qquad \Rightarrow \qquad [r] = [m_2] = T^{-1}$$

   En adoptant les variables sans dimension x, y et τ, les équations deviennent

$$\begin{cases} \dfrac{jr}{m_1}\dfrac{dx}{d\tau} = j - jx \\ r\kappa\dfrac{dy}{d\tau} = \dfrac{rj}{m_1}x(1-y) - m_2\kappa y \end{cases}$$

soit, après simplification,

$$\begin{cases} \dfrac{dx}{d\tau} = \dfrac{m_1}{r}(1-x) \\ \dfrac{dy}{d\tau} = \dfrac{j}{\kappa m_1} x(1-y) - \dfrac{m_2}{r} y \end{cases}$$

qui correspond à la forme (†) attendu à condition de poser

$$\alpha = \frac{m_1}{r}, \qquad \beta = \frac{j}{\kappa m_1} \qquad \text{et} \qquad \gamma = \frac{m_2}{r}$$

iii. *En considérant la forme adimensionnelle (†) des équations, déterminez les éventuelles configurations d'équilibre du système et leur stabilité.*

Les configurations d'équilibre sont les solutions $(x^\star, y^\star)$ de

$$\begin{cases} 0 = \alpha(1-x^\star) \\ 0 = \beta x^\star (1-y^\star) - \gamma y^\star \end{cases}$$

soit

$$(x^\star, y^\star) = \left(1, \ \frac{\beta}{\beta+\gamma}\right)$$

Pour étudier la stabilité, on pose

$$x = x^\star + \varepsilon = 1 + \varepsilon, \qquad y = y^\star + \eta = \frac{\beta}{\beta+\gamma} + \eta$$

et on linéarise les équations pour les perturbations (ε, η) obtenues à partir de (†), *i.e.*

$$\begin{cases} \dfrac{d\varepsilon}{d\tau} = -\alpha\varepsilon \\ \dfrac{d\eta}{d\tau} = \beta(1+\varepsilon)\left(1 - \dfrac{\beta}{\beta+\gamma} - \eta\right) - \gamma\left(\dfrac{\beta}{\beta+\gamma} + \eta\right) \sim \dfrac{\beta\gamma}{\beta+\gamma}\varepsilon - (\beta+\gamma)\eta \end{cases}$$

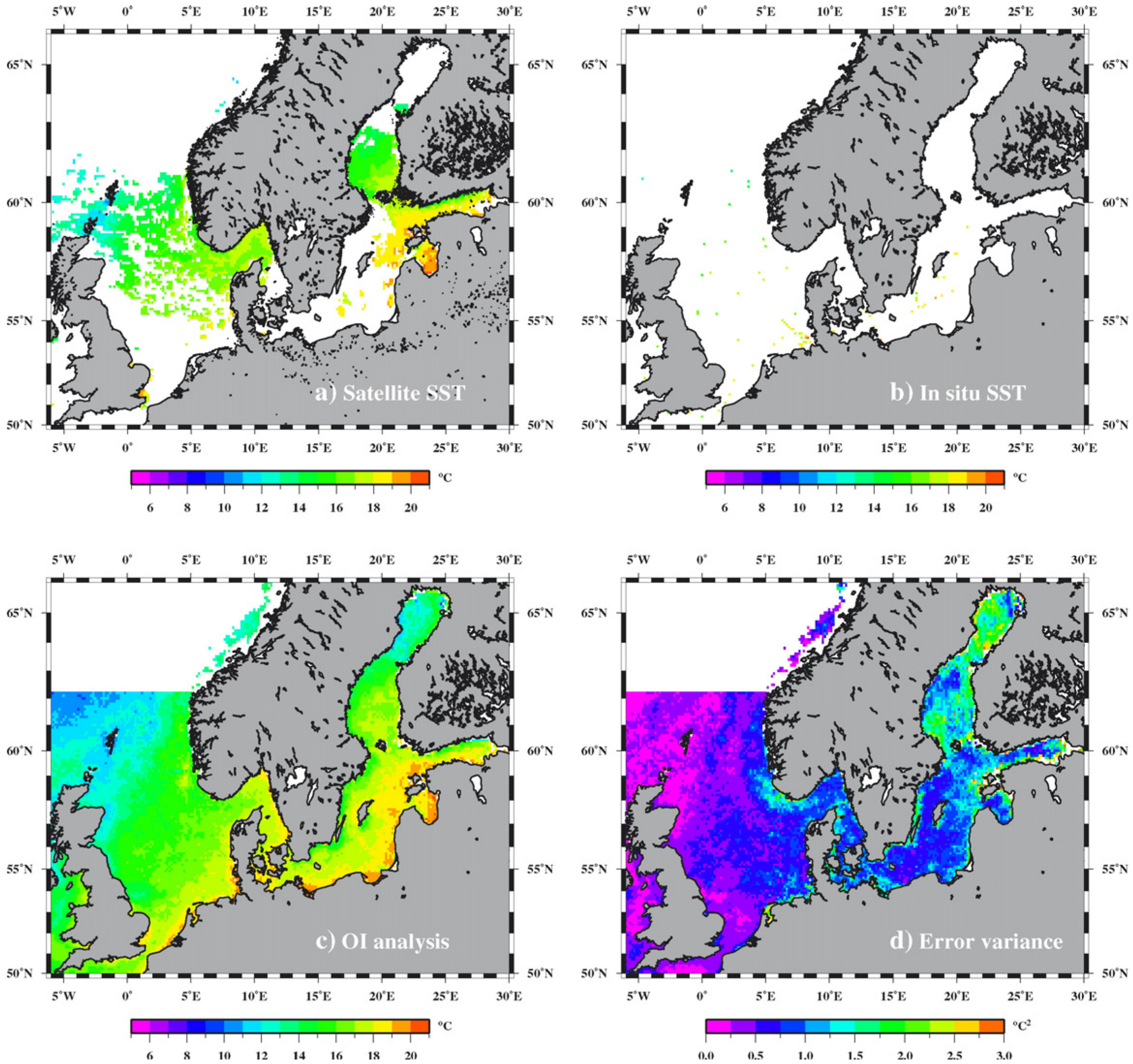
Les équations peuvent être écrites sous la forme matricielle

$$\frac{d}{d\tau}\begin{pmatrix} \varepsilon \\ \eta \end{pmatrix} = \begin{pmatrix} -\alpha & 0 \\ \dfrac{\beta\gamma}{\beta+\gamma} & -(\beta+\gamma) \end{pmatrix} \begin{pmatrix} \varepsilon \\ \eta \end{pmatrix}$$

Les valeurs propres de la matrice de communauté sont $-\alpha$ et $-(\beta+\gamma)$. Celles-ci étant toutes deux strictement négatives, les perturbations décroissent exponentiellement au cours du temps et l'équilibre est stable.

Question III

La figure ci-dessous est extraite de l'article 'Optimal interpolation of sea surface temperature for the North Sea and Baltic Sea' publié par Hoyer et She en 2005 dans le Journal of Marine Systems. Elle montre les données utilisées pour générer la carte de la température (données satellitaires et mesures in situ) et le résultat de l'analyse optimale (OI) de ces données.



i. *Expliquez les difficultés pour combiner ainsi des données expérimentales d'origines et de distributions spatiales différentes.*

Les données satellitaires et in situ sont entâchées d'erreurs de mesures différentes ; il convient de pondérer de façon appropriée les données en fonction de ces erreurs de mesures, *i.e.* en donnant plus de poids aux données les plus précises.

La distribution non homogène des données doit être compensée pour éviter l'effet de 'clustering'. Les données mesurées en des points proches l'un de l'autre portent des informations semblables ou partiellement redondantes. D'une part, les poids attribués à ces points doivent être réduits par rapport à des mesures indépendantes. D'autre part, la proximité entre les points de mesure peut être exploitée pour réduire l'incertitude sur les mesures.

ii. *Expliquez le principe de l'analyse objective. En quoi l'analyse est-elle objective ?*

L'analyse objective permet de générer un champ interpolé en combinant les données mesurées avec un champ de référence, correspondant à l'information disponible a priori avant la prise en compte des mesures.

Le champ analysé x_a peut être décrite par

$$x_a = x_{ref} + W(y_{obs} - y_{ref})$$

où y_{obs} désigne le vecteur des observations et où x_{ref} et y_{ref} sont les valeurs du champ de référence respectivement aux points où l'analyse est effectuée et aux points de mesure. La matrice W contient les poids respectifs des différentes observations qui sont donc combinées linéairement.

L'erreur de l'analyse peut être décrite par le vecteur $\varepsilon_a = \mathsf{x}_a - \mathsf{x}_r$ où x_r désigne le champ réel.

L'analyse est dite objective car les poids W sont fixés par application d'un critère objectif d'optimalité : W est la matrice qui réalise le minimum de la variance des erreurs $E[\varepsilon_a \varepsilon_a^{\mathrm{T}}]$ (où $E[\cdot]$ désigne l'espérance mathématique).

Le calcul de W prend en compte la variance/co-variance du champ de référence — donc la variabilité du champ réel — ainsi que la variance/co-variance des observations — donc les erreurs de mesure— au sujet desquels il convient de faire des hypothèses ou de disposer d'informations statistiques appropriées. En particulier, le calcul tient compte de la correlation spatiale du champ étudié et des observations.

iii. *Quelle est la signification de 'l'error variance' représentée dans la dernière figure ?*

L'error variance' mesure l'erreur résiduelle du champ résultant de l'analyse objective. La variance est égale à la variance du champ de référence réduite par la prise en compte des mesures observées. Celle-ci est donc réduite dans les zones où des données (précises) sont disponibles et importante là où les observations n'apportent pas ou peu d'information.

Là où la variance de l'erreur est grande, par exemple ici dans la partie nord de la mer Baltique où peu de données sont disponibles, l'incertitude sur l'analyse objective est importante.

Question IV

On souhaite interpoler les données (x_i, T_i) $(i = 1, 2, \ldots N)$ obtenues en mesurant la température T_i en différents points x_i le long d'un transect rectiligne. Pour chacune des méthodes d'interpolation envisagées, déterminez la liste des propriétés du champ interpolé correspondant. Au besoin, nuancez votre réponse.

	Interpolation linéaire par morceau	*Interpolation polynômiale (Lagrange)*	*Interpolation spline*	*Interpolation par distance inverse*	*Krigeage ('ordinary krigging')*	*Analyse objective*
Le champ interpolé varie continûment	X	X	X	X	X	X
La dérivée du champ interpolé varie continûment.	-	X	X	X[1]	X[1]	X[1]
Les valeurs interpolées appartiennent au même intervalle de variation que les valeurs mesurées.[2]	X	-	-	X	-	-
L'interpolation passe exactement par les points de mesure.	X	X	X	X	X	-
L'interpolation est obtenue par minimisation de la courbure globale du champ.	-	-	X	-	-	-
L'interpolation tient compte de l'effet de 'clustering' des données. -	-	-	-	-	X	X
L'ajout de nouvelles données modifie seulement localement l'interpolation.	X	-	-	-[3]	-	-

[1] Les dérivées sont discontinues si les distances sont pondérées par un exposant $p \leq 1$.

[2] Dans le cas d'une méthode d'interpolation dans laquelle les données apparaissent linéairement, *i.e.* selon un expression du type

$$x^{interpol} = \sum_i w_i x_i^{obs}$$

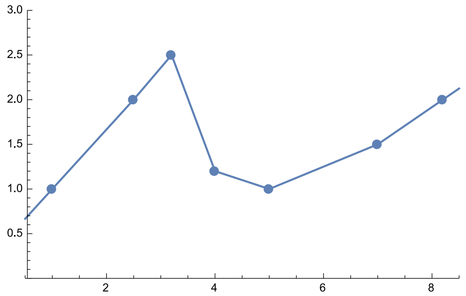
les valeurs interpolées appartiennent au même intervalle de variation que les valeurs mesurées si et seulement si les poids w_i sont tous positifs et leur somme est égale à l'unité. C'est le cas dans l'interpolation linéaire par morceau ou dans l'interpolation par distance inverse. Dans le krigeage, les poids peuvent être négatifs et la propriété n'est donc pas vérifiée.

[3] La distance à laquelle les données influencent le champ interpolé peut être librement limitée en ne prenant en compte dans la détermination du champ interpolé en un point que les mesures effectuées à une distance d de ce point inférieure à une certaine valeur critique d_{cr}.

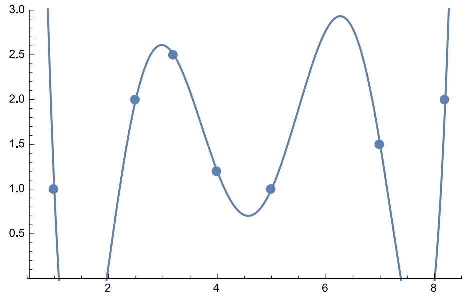
Fondamentalement, l'analyse optimale et le krigeage ne diffèrent que par la prise en compte ou non d'un champ de référence. Les deux méthodes donnent donc des résultats qui possèdent les mêmes propriétés de régularité du champ. La différence essentielle réside dans la pondération entre les observations et le champ de référence introduite par l'analyse optimale ; le champ analysé ne passe donc pas exactement par les valeurs mesurées si celles-ci sont supposées entachées d'une erreur. Le champ obtenu par krigeage passe exactement par les points de mesure.

À titre d'illustration, les figures ci-dessous présentent le résultat de l'interpolation spatiale unidimensionnelle de données en utilisant les différentes méthodes considérées.

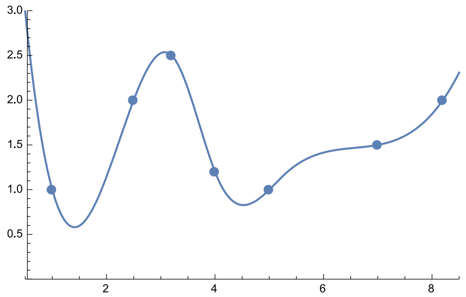
Interpolation linéaire par morceau



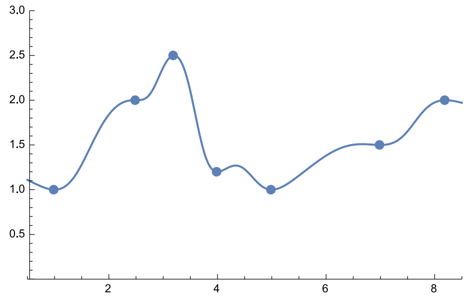
Interpolation polynomiale



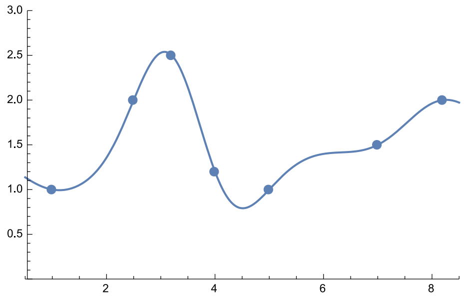
Interpolation spline



Interpolation par distance inverse



Krigeage



Analyse objective