I. méthodes d'interpolation spatiale

La variation spatiale et spatio-temporelle des phénomènes physiques et socio-économiques, n'est connue qu'a travers des d'observations géoréférencées, pas nécessairement réparties régulièrement. On cherche alors à estimer les valeurs prises par ces phénomènes observés en d'autres points de l'espace. On parle alors d'estimation spatiale : c'est une procédure consistant à estimer la valeur d'une grandeur en un site à partir d'échantillons de cette grandeur récoltés dans d'autres sites.

Ce besoin s'applique à de nombreux domaines où la connaissance de la distribution spatiale de phénomènes est importante : altimétrie, gravimétrie, météorologie, géologie, minéralogie, etc.

Dans cette partie, nous allons donc étudier les méthodes permettant l'estimation et l'interpolation de données géoréférencées. Cette partie c'est une introduction à certaines méthodes existantes utilisées pour la prévision d'une valeur inconnue à partir d'observations. On parle alors d'interpolation pour l'estimation de cette valeur.

Les méthodes de prévision, se divisent usuellement en deux groupes, selon les modèles mathématiques sur lesquels elles reposent :

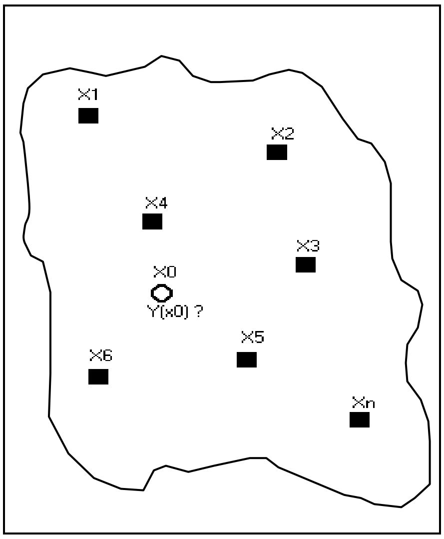
- **Méthodes déterministes :** ces méthodes se basent sur des propriétés purement mathématiques, généralement géométriques, sans tenir compte du phénomène physique considéré.

- **Méthodes stochastiques :** elles supposent une modélisation probabiliste du phénomène, dont les observations résultent de la réalisation de variables aléatoires; ces méthodes font alors appel à des modèles découlant de l'analyse statistique des données considérées. On parle alors de techniques géostatistiques bas é sur la théorie des variables régionalisées.

I.1 Méthodes d'interpolation déterministes

I.1.1 Introduction

Soit une propriété Z qui varie dans l'espace géographique, connue de façon discontinue en n points d'observation localisés par leurs coordonnées géographiques x_i.



Exemples De Telles Proprietes Sont :

- Hauteur De Pluies Mesurees Dans Quelques Stations Meteorologiques,
- Profondeur Du Sol Mesuree A La Tariere,
- Teneurs En Metaux Lourds Dans Le Sol,
- Teneur En Or Dans Des Couches Sedimentaires A Partir De Quelques Forages...

Le problème peut être formalisé simplement :

Soit x_0 un point quelconque de cet espace. Peut on estimer la valeur de Z en x_0 à partir des valeurs Z(xi) connues ? ; Sachant faire cela, peut-on obtenir la carte des variations de la propriété dans l'espace géographique?

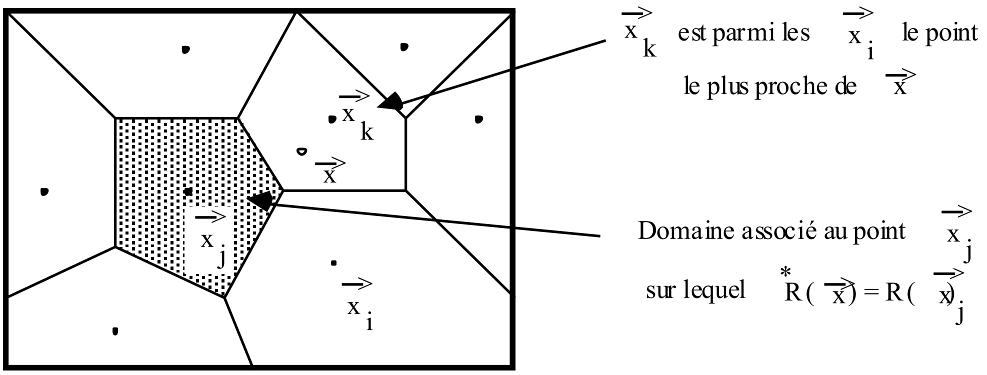
Deux méthodes intuitives vont nous aider à introduire les problèmes que nous aurons à résoudre.

I.1.2 La méthode des polygones de Thiessen

Pour estimer la valeur en un point quelconque, on peut prendre la valeur du point observé le plus proche. Cette méthode est connue sous les titres de «Méthode des polygones de Thiessen» ou de « Proximal approximation ».

Dès 1911, Thiessen proposait une méthode pour évaluer des moyennes spatiales sur un domaine à partir d'informations ponctuelles. D'une façon sous-jacente, la « méthode de Thiessen » est une technique d'interpolation basée sur « la loi du plus proche voisin ». Supposons qu'une variable $Z(\vec{x})$ soit connue en un certain nombre de points notés $\vec{x}_i$. Thiessen admet qu'en un point quelconque $\vec{x}$ différent des $\vec{x}_i$, on peut évaluer $Z(\vec{x})$ vrai inconnu par l'estimateur $Z^*(x)$ ainsi défini : $Z^*(\vec{x}) = Z(\vec{x}_k)$ ($\vec{x}_k$ étant parmi tous les $\vec{x}_i$, celui qui est le plus proche de $\vec{x}_i$)

Comme l'illustre le schéma ci dessous, la méthode de Thiessen revient à interpoler la fonction $Z(\vec{x})$ par une fonction $Z^*(x)$ en escalier, constante par polygone. Les polygones, appelés polygones de Thiessen, sont construits à partir des médiatrices des segments joignant les points d'échantillonnage $\vec{x}_i$.



Manifestement, l'unique avantage de cette méthode réside dans sa simplicité. A priori, on ne dispose d'aucune information objective sur la représentativité des interpolations. Le problème est réside dans le nombre d'échantillonnage ; si la densité des points d'échantillonnage $\vec{x}_i$ est "très forte", la représentativité peut être à peu près correcte dans d'autres cas, au contraire les résultats peuvent être très décevants. Mais, cette approche est limitée car :

- l'information utilisée pour estimer la valeur en un point est faible, elle se limite à la prise en compte d'un seul point observé.

- elle suppose sans le vérifier que le point estimé est corrélé au point observé. Or, si la variable est distribuée de façon aléatoire dans l'espace, cette hypothèse est invalide.

- il est impossible d'avoir une idée de la précision des estimations obtenues.

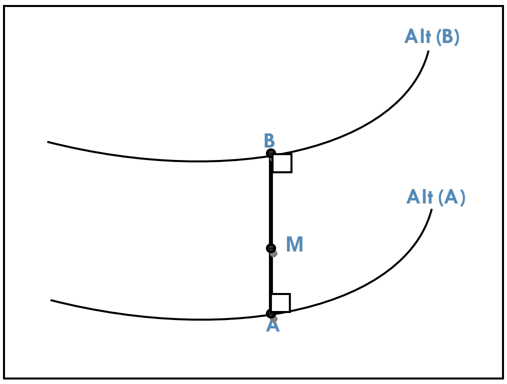
On se limitera donc à l'emploi de telles méthodes pour avoir une idée de la structure spatiale générale avant de passer à des méthodes plus détaillées.

II.1.3 Moyenne pondérée des observations les plus proches

Cette méthode est également intuitive et peut être illustrée par la lecture d'une carte topographique.

Connaissant l'altitude en deux courbes de niveaux, une estimation simple de l'altitude en un point quelconque M peut être donnée par l'expression :

$$\mathbf{Alt\ (M) = Alt(A) + \frac{|AM|}{|AB|}.\left(Alt(B) - Alt(A)\right)}$$



Dans ce cas, on accorde, dans l'estimation du point M, un poids d'autant plus fort aux points observés A ou B qu'ils sont proches de M.

Cette méthode dite "d'interpolation linéaire" utilise plus d'information que celle des polygones de Thiessen, mais les limites citées précédemment sont aussi valables pour celle ci :

- On suppose que la variation de la propriété d'un point observé à l'autre est linéaire sans le vérifier;

- On n'a pas d'idées de la précision des estimations faites.

Conclusion

Ces deux exemples montrent bien que l'estimation d'une propriété dans l'espace géographique suppose deux étapes:

1. une phase d'analyse de la structure spatiale de la propriété étudiée, pour savoir comment sont corrélés entre eux les points observés (problème d'hypothèse de linéarité).
2. une phase d'estimation proprement dite tenant compte de la structure spatiale précédemment identifiée (utilisation de la fonction de structuration ainsi définie pour l'estimation de la variable).