Club Math

Lundi 5 février 1996 à 17 h

Visualiser des données 3D

Bernard Vuilleumier

Centre informatique
pédagogique (CIP)
Rue Théodore-de-Bèze 2
Case Postale 3144
1211 GENÈVE 3
Tél: (022) 318.05.30
Fax: (022) 318.05.35
Directeur:
Raymond Morel

Lettre n° 46

Créer des images à partir de données numériques est un problème qui revêt une grande importance pour tous ceux qui utilisent un ordinateur. Les logiciels disponibles pour réaliser cette tâche sont de plus en plus nombreux et certains offrent de remarquables possibilités. Pour parler de visualisation de données numériques, il est nécessaire de développer un système de classification. Une caractéristique de base d'un *ensemble de données* est sa *dimension*, c'est-à-dire le nombre d'item que comprend chaque donnée. Par exemple un ensemble de points sur une droite a la dimension 1, un ensemble de points sur un plan la dimension 2 et un ensemble de points dans l'espace la dimension 3. Une autre caractéristique est le nombre de *variables indépendantes* et le nombre de *variables dépendantes* d'une donnée. Si chaque donnée ne contient qu'une variable dépendante, c'est une donnée scalaire. Lorsqu'il y a plus d'une variable dépendante, c'est une donnée vectorielle. Le nombre d'entrées représentant les variables indépendantes peut être utilisé pour déterminer la structure graphique appropriée pour la visualisation des données. Les ensembles de données de dimension 3 par exemple qui possèdent une variable indépendante sont bien représentés par une courbe, alors qu'une surface est plus appropriée pour ceux à deux variables indépendantes.

Pour obtenir une fonction mathématique à partir de données numériques, deux approches sont possibles: interpoler et ajuster

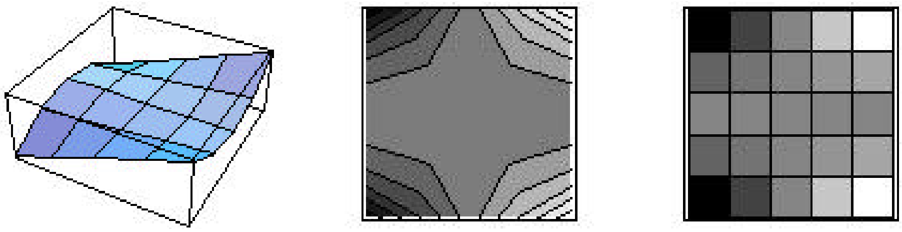

Fig. 1: Différentes représentations d'un ensemble de données à 3 dimensions et à deux variables indépendantes (données scalaires). A gauche les données sont représentées par une surface, au centre par des lignes de niveau et à droite par des densités de gris.

Lorsqu'on construit une fonction à partir de données numériques, on est amené a prendre des décisions

Pour obtenir une fonction mathématique à partir de données numériques, deux approches sont possibles: interpoler et ajuster (to fit en anglais). L'*interpolation* s'applique lorsque les données sont précises, lorsqu'aucune donnée ne comporte d'incertitude significative. Dans ce cas, on cherche à relier les points en supposant qu'ils sont tous bons. Le fit est approprié lorsque les données sont imprécises et lorsque certains points doivent être éliminés. Il consiste à ajuster au mieux – en minimisant une grandeur – une ligne aux points. Mais pour réaliser un fit, la forme de la fonction doit être connue. On commence en général par représenter les données en utilisant une technique d'interpolation. Un examen visuel permet de détecter d'éventuelles données erronées et suggère une forme de fonction pour réaliser le fit.

Quand les données sont régulièrement espacées et ordonnées rectangulairement, on peut facilement les représenter à l'aide des fonctions standard des logiciels

Lorsque des données $\{\{x_1, y_1, z_1\}, \{x_2, y_2, z_2\}, \ldots\}$ de dimension 3 comportent deux variables indépendantes x et y, on les représente en général par une surface. La hauteur de la surface donne la variation de la variable dépendante z en fonction des deux variables indépendantes x et y. Un graphique de ce type donne le lieu des points $\{x, y, f(x, y)\}$. Pour dessiner cette surface, il faut décider comment relier les points. Si les données sont régulièrement espacées, il suffit de les ordonner rectangulairement en les mettant sous la forme $\{\{z_{11}, z_{12}, \ldots\}, \{z_{21}, z_{22}, \ldots\}, \ldots\}$ où chaque z_{ij} est un nombre donnant la hauteur associée aux i^e et j^e variables indépendantes et d'utiliser les fonctions standard des logiciels. Si les données ne sont pas réparties régulièrement sur une grille, on peut effectuer une triangulation de telle sorte que chaque point soit le sommet d'au moins un triangle et que chaque triangle soit le plus proche possible d'un triangle équilatéral.

Mais si elles ne sont pas réparties régulièrement sur une grille, il faut trouver un moyen de relier les points puis de les représenter

Travaux pratiques

Mots clefs
ensemble de données, dimension, variable indépendante, variable dépendante, interpolation.

Pour se familiariser avec la notion d'ensemble de données ordonnées rectangulairement

Exercice 1
Soit l'ensemble de données suivant:
{{-2, -2, -8}, {0, -2, 0}, {1, 0, 0}, {-2, 0, 0}, {2, -2, -8}, {2, 2, 8}, {1, 2, 2}, {1, -1, -1}, {-1, -2, -2}, {-1, 1, 1}, {0, 0, 0}, {0, -1, 0}, {1, -2, -2}, {2, 1, 4}, {-2, 2, 8}, {-1, -1, -1}, {-2, -1, -4}, {-1, 0, 0}, {0, 1, 0}, {-2, 1, 4}, {2, -1, -4}, {1, 1, 1}, {-1, 2, 2}, {0, 2, 0}, {2, 0, 0}}
a) Quelle est sa dimension?
b) Représentez ces données de différentes manières en supposant que les deux premiers item de chacune d'elles sont des variables indépendantes.

Pour découvrir une technique d'organisation de données qui ne sont pas réparties régulièrement sur une grille

Exercice 2 (pour répondre à une question de mon collègue et ami Claude Renaud)
Les données recueillies par l'ingénieur ou le physicien ne sont en général ni organisées rectangulairement ni régulièrement espacées et ne peuvent pas être mises sous cette forme. Comment alors construire une surface à partir d'un ensemble tridimensionnel de données non organisées? Une possibilité consiste à grouper les variables indépendantes en entités permettant de construire une surface. Une entité possible est le triangle. Le problème peut être résolu en formant des triangles dans le plan *xy* à l'aide d'une triangulation de Delaunay. Cette dernière consiste à construire des triangles à partir d'un ensemble de points de telle sorte que chaque point soit le sommet d'au moins un triangle et que le plus petit angle de chaque triangle soit maximum. Chaque triangle est alors aussi proche que possible d'un triangle équilatéral (il existe un module *Mathematica* effectuant des triangulations de Delaunay).
a) Créez un ensemble bidimensionnel de 20 données aléatoires.
b) Effectuez une triangulation de Delaunay de ces données et représentez-là.

Pour appliquer une triangulation de Delaunay à un ensemble de données

Exercice 3
Soit l'ensemble tridimensionnel de données:
data3D = Table[{x = Random [ ], y = Random [ ], z = x y},{5}].
a) Projetez ces données dans le plan *xy* puis effectuez un triangulation de Delaunay.
b) Représentez les données à l'aide d'une surface.

Pour illustrer d'intéressantes possibilités de représentation et d'interpolation à l'aide de modules Mathematica

Exercice 4
En utilisant les commandes du module **SurfaceGraphics3D** et **TriangularInterpolate**:
a) Représentez l'ensemble suivant: data3D = Table[{x = Random[Real, {-2, 2}], y = Random[Real, {-2, 2}], Exp[- Sqrt[x^2 + y^2]]}, {250}].
b) Appliquez une méthode d'interpolation à la surface obtenue sous a).
N.B. Les modules **SurfaceGraphics3D** et **TriangularInterpolate** ne sont pas fournis avec *Mathematica* et ne font pas partie du domaine public. Ils sont décrits dans l'ouvrage de Tom Wickham-Jones, *Mathematica Graphics*, Springer-Verlag, New-York: 1994. Ils se trouvent sur la disquette qui accompagne le livre.

Prochaine réunion: lundi 4 mars 1996 à 17 h.