

2. Übungsblatt zur Vorlesung Algorithmen für geographische Informationssysteme (WS 2024/25)

Aufgabe 1 – Methode der kleinsten Quadrate

In dieser Übung implementieren Sie die Methode der linearen kleinsten Quadrate, wie sie in den Vorlesungen beschrieben wurde. Insbesondere soll ihr Programm allgemeine Instanzen des Höhenmessproblems („Beispiel 2“) lösen. Verwenden Sie eine imperative Programmiersprache wie C(++/#), Java oder Python, nicht ein Computer-Algebrasystem wie Mathematica. Für die Matrizenoperationen (Multiplikationen, Inversionen, Lösung eines Systems $Ax = b$, ...) können Sie Bibliotheken wie Jama oder Eigen verwenden, jedoch keine Methoden, die direkt eine „Lösung der kleinsten Quadrate“ liefern.

Wir interessieren uns für die Höhen von n Punkten: einem Punkt P_1 und weiteren $n - 1$ Punkten. Wir definieren das Koordinatensystem, indem wir die Höhe von P_1 auf 0 setzen, das heißt $h_1 = 0$. Nun möchten wir die anderen Höhen basierend auf einer Liste von Höhendifferenzmessungen schätzen. Wir nehmen an, dass die Messungen Fehler enthalten können, diese Fehler jedoch statistisch unabhängig sind.

Eingabe

Ihr Programm sollte die folgende Eingabe (über den Standard-Stream wie `cin`, `System.in` oder `input`) akzeptieren. Auf der ersten Zeile steht eine einzelne ganze Zahl $n > 1$, die Anzahl der Punkte. In der zweiten Zeile steht eine einzelne ganze Zahl $m \geq 1$, die Anzahl der Messungen. Danach folgen m Zeilen, von denen jede die folgenden vier Zahlen enthält, getrennt durch Leerzeichen:

- Den Index i des Punktes, *von dem* die Messung stammt.
- Den Index j des Punktes, *zu dem* die Messung durchgeführt wurde.
- Den Wert der Messung, d. h. eine Messung von $h_j - h_i$.

Das Beispiel aus den Vorlesungsfolien würde wie folgt dargestellt:

4		
5		
1	2	4.1
2	3	-7
3	4	1.1
4	1	1.2
4	2	5.4

- a) Ihr Programm sollte Folgendes ausgeben: Eine Schätzung der Höhen $h_2, \dots, h_n$ nach der Methode der kleinsten Quadrate sowie die „korrigierten Messungen“.
- b) Wir möchten nun Genauigkeiten der Messungen bestimmen.

Starten Sie mit einem bekannten Satz von Höhen. Dies impliziert eine vollständige, konsistente Menge von Messungen. Ihr Programm aus Teil (a) sollte die korrekten Höhen daraus berechnen können.

Fügen Sie automatisch Rauschen zu den Messungen hinzu (z. B. einen unabhängig normalverteilten Fehler für jede Messung). Wie gut kann Ihr Programm die Höhen nun schätzen?

- c) Wir können die Basisvarianzen des Rauschens wie folgt schätzen. Bei n Messungen und u Unbekannten gibt $\sigma_0^2 = \frac{v^T v}{n-u}$ die *Basisvarianz* an. Die Varianzen der einzelnen Höhen erlangt man von der Diagonalen der Matrix $\sigma_0^2 \cdot (A^T A)^{-1}$.

Für das Beispiel aus der Vorlesung erhalten wir

$$v^T v = (0.1 \quad 0.2 \quad 0.2 \quad 0.1 \quad 0.1) \cdot \begin{pmatrix} 0.1 \\ 0.2 \\ 0.2 \\ 0.1 \\ 0.1 \end{pmatrix} = 3 \cdot 0.1^2 + 2 \cdot 0.2^2 = 0.11 \text{m}^2$$

$$\sigma_0^2 = \frac{v^T v}{n-u} = \frac{0.11 \text{m}^2}{5-3} = 0.055 \text{m}^2$$

$$\sigma_0^2 \cdot (A^T A)^{-1} = 0.055 \text{m}^2 \cdot \begin{pmatrix} 0.625 & 0.500 & 0.375 \\ 0.500 & 1.000 & 0.500 \\ 0.375 & 0.500 & 0.625 \end{pmatrix} = \begin{pmatrix} 0.0344 & 0.0275 & 0.0206 \\ 0.0275 & 0.0550 & 0.0275 \\ 0.0206 & 0.0275 & 0.0344 \end{pmatrix} \text{m}^2.$$

Und damit

$$\sigma_{h_2} = \sigma_{h_4} = \sqrt{0.0344 \text{m}^2} = 0.185 \text{m}$$

$$\sigma_{h_3} = \sqrt{0.0550 \text{m}^2} = 0.235 \text{m}$$

.

Wie gut schätzt das Programm die Basisvarianz Ihrer Messungen? (Sie kennen die Varianz, da Sie das Rauschen selbst hinzugefügt haben.)

Wie beeinflussen die Varianzen und die Menge der Messungen die Qualität der Lösung?

Aufgabe 2 – Kartenbeschriftungen

Für jeden Punkt einer gegebenen Menge P sei dessen rechteckiges Label gegeben. Sei nun $ts(P)$ die maximale Anzahl der gelabelten Punkte im *top slider model*. Analog seien $op(P)$ und $tp(P)$ für das *one position* bzw. *two position model* definiert (jeweils auf der oberen Seite). Wir gehen davon aus, dass alle Rechtecke Höhe 1 haben.

a) Wie hoch kann das Verhältnis $\frac{ts(P)}{tp(P)}$ werden, wenn alle Rechtecke

- Quadrate sind?
- gleich lange Rechtecke sind?
- beliebige Rechtecke sind?

b) Wie hoch kann das Verhältnis $\frac{tp(P)}{op(P)}$ werden, wenn alle Rechtecke

- Quadrate sind?
- gleich lange Rechtecke sind?
- beliebige Rechtecke sind?