Technische Universität München
Institut für Informatik
Prof. Dr. Thomas Huckle
Dipl.-Tech. Math. Stefanie Schraufstetter
Michael Lieb, M. Sc.

SS 2010

Numerisches Programmieren, Übungen

4. Übungsblatt: Lineares Ausgleichsproblem, Regularisierung

1) Lineares Ausgleichsproblem

(modifizierte Aufgabe der Semestralklausur des WS 2007/08)

Bei der linearen Ausgleichsrechnung geht es darum, für einen Satz von k Punkten P_i ($i = 1, \ldots, k$) eine Ausgleichsgerade zu finden, die zu allen Punkten einen möglichst geringen Abstand hat. Dies kann man erreichen, indem man folgendes Minimierungsproblem löst:

$$\min_z ||b - Az||_2. \tag{1}$$

Im Folgenden werden wir ein Minimierungsproblem für den zweidimensionalen Fall betrachten, d.h. $z \in \mathbb{R}^2$. Um die Matrix A (Dimension $k \times 2$) und den Vektor b (Länge k) zu erhalten, stellt man ein überbestimmtes lineares Gleichungssystem auf, indem man für jeden Punkt P_i ($i = 1, \ldots, k$) die allgemeine Geradengleichung $y(x) = m \cdot x + t$ ansetzt. Die Unbekannten sind dann die Steigung m und der Achsenabschnitt t der Ausgleichsgeraden, die in dem Vektor $z = (t, m)^T$ zusammengefasst werden.

Wenn man die Matrix A mit Hilfe orthogonaler Transformationen auf rechte obere Dreiecksgestalt bringt,

$$Q^T A = \left[\begin{array}{ccc} * & \dots & * \\ & \ddots & \vdots \\ \mathbf{0} & & * \\ \hline & \mathbf{0} & \end{array}\right] = \left[\begin{array}{c} R \\ \hline \mathbf{0} \end{array}\right], \tag{2}$$

ist die Lösung des Minimierungsproblems (1) äquivalent zur Lösung des Gleichungssystems

$$Rz = \beta_1, \quad \text{wobei } Q^T b = \left(\begin{array}{c} \beta_1 \\ \hline \beta_2 \end{array}\right). \tag{3}$$

Für die Transformation der Matrix A verwenden wir 3D-Givens-Rotationen. Dabei dreht man mittels Elementardrehungen $G_{i,j}$ jeweils in einer x_i-x_j-Ebene, um den Eintrag $A_{i,j}$ zu Null zu machen. Ziel ist eine rechte obere Dreiecksmatrix $R \in \mathbb{R}^{k\times 2}$, die aus Anwendung von Kombinationen von $G_{i,j}$ auf A hervorgeht.

Die Elementardrehungen $G_{i,j}$ sind dabei so definiert, dass sie lediglich jeweils die i-te und j-te Zeile modifizieren:

$$G_{i,j} = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & c & & s & & \\ & & & \ddots & & & \\ & & s & & -c & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}. \tag{4}$$

So ergeben sich zum Beispiel im Dreidimensionalen ($k = 3$) die folgenden drei Elementardrehungen $G_{i,j}$, bei deren Anwendung jeweils die erste, zweite bzw. dritte Zeile unverändert bleibt:

$$G_{3,2} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c & s \\ 0 & -s & c \end{pmatrix}, \quad G_{3,1} = \begin{pmatrix} c & 0 & s \\ 0 & 1 & 0 \\ -s & 0 & c \end{pmatrix}, \quad G_{2,1} = \begin{pmatrix} c & s & 0 \\ -s & c & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{5}$$

Welche der Strukturen zu wählen ist, entscheidet sich danach, welcher Eintrag zu Null gedreht werden soll. Der zu annullierende Eintrag $A_{i,j}$ sowie das zugehörige Diagonalelement $A_{j,j}$ der betrachteten Spalte dürfen sich ändern, die restlichen Einträge der jeweiligen Matrix-Spalte müssen gleich bleiben.
Für die Einträge c und s in den Drehungen gilt analog zum 2D-Fall folgender Zusammenhang (mit $d := A_{j,j}$ = Diagonalelement und $a := A_{i,j}$ = zu annullierendes Element der aktuellen Spalte):

$$\begin{aligned} c &= \cos(\varphi) = \frac{d}{\sqrt{d^2 + a^2}} \\ s &= \sin(\varphi) = \frac{a}{\sqrt{d^2 + a^2}} \end{aligned}$$

Nachfolgend soll nun die Ausgleichsgerade für die drei folgenden Punkte ($k = 3$) bestimmt werden:

$$P_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \quad P_2 = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad P_3 = \begin{pmatrix} x_3 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

i) Tragen Sie in ein Koordinatensystem die gegebenen Punkte ein. Stellen Sie die zugehörige Matrix $A \in \mathbb{R}^{3\times 2}$ sowie den Vektor $b \in \mathbb{R}^3$ auf. Beachten Sie dabei die Reihenfolge der Unbekannten, die über $z = (t, m)^T$ festgelegt ist.

ii) Berechnen Sie die rechte obere Dreiecksmatrix R mit Hilfe passender Elementardrehungen $G_{i,j}$ nach (5). Es ist vorteilhaft, mit dem Element $A_{2,1}$ zu beginnen.

iii) Berechnen Sie die Lösung $z = (t, m)^T$ des Minimierungsproblems (1) mit Hilfe der QR-Zerlegung aus (ii). Es reicht dabei, die Elementardrehungen $G_{i,j}$ zu nutzen, ohne eine neue Matrix explizit aufzustellen. Zeichnen Sie die resultierende Ausgleichsgerade in das Bild aus Teilaufgabe (i) ein.

iv) Zeigen Sie, dass die Lösung des Systems $Rz = \beta_1$ basierend auf orthogonalen Transformationen Q^T tatsächlich äquivalent zur Lösung der Minimierungsaufgabe $\min_z ||b - Az||_2$ ist.

2) Computertomographie

Bei der Computertomographie geht es darum, Informationen über das Innere eines Gegenstandes zu erhalten, ohne den Gegenstand zu zerlegen. Dabei soll insbesondere die Dichte des Gegenstandes bestimmt werden, um daraus die einzelnen Bestandteile wie Knochen oder Organe zu identifizieren. Dazu schickt man Strahlen durch den Körper und misst den Identitätsverlust der Strahlen, welcher aus dem Durchdringen des Körpers resultiert.

Wir wollen hier das Verfahren der Computertomographie an einem vereinfachten Beispiel betrachten. Dazu sei eine Fläche $Q = [0, 3]^2$ gegeben, die in 3×3 Quadrate Q_k gleicher Größe mit jeweils konstanten Dichten ρ_k eingeteilt ist (vgl. Abb. 1). Schickt man Strahlen S_j $(j = 1, \ldots, n)$ durch den Gegenstand, so gilt für den Intensitätsverlust v_k beim Durchdringen der Fläche, die Gleichung

$$v_j = \sum_k L_{j,k} \cdot \rho_k, \tag{6}$$

wobei ρ_k jeweils die Dichte des Quadrats Q_k bezeichnet und $L_{j,k}$ die Länge des Strahls S_j durch das Quadrat Q_k. Um die Dichten ρ_k zu bestimmen, löst man anschließend das lineare Ausgleichungsproblem

$$\min_r ||Lr - v||_2 \tag{7}$$

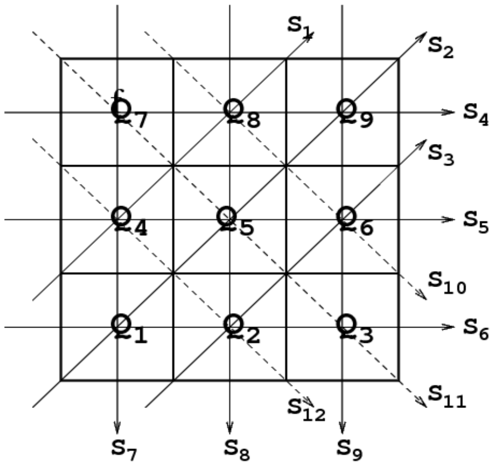

Abbildung 1: In Zellen eingeteilte Fläche mit CT-Strahlen

durch Lösen der Normalengleichung

$$L^T L r = L^T v, \tag{8}$$

wobei die Bezeichnungen $L = (L_{j,k})_{j,k} \in \mathbb{R}^{n \times m},\ r = (\rho_k)_k \in \mathbb{R}^m,\ v = (v_j)_j \in \mathbb{R}^n$ gelten.

i) Zunächst benutzen wir für die Messungen die 9 Strahlen $S_1, \dots S_9$ (vgl. Abb. 1). Stellen Sie hierfür die Matrix L des Gleichungssystems (6) auf. Sind die Matrizen L bzw. $L^T L$ regulär? Was bedeutet dies?

ii) Welche Intensitätsverluste misst man, wenn alle Zellen Q_j mit ungeradem Index k die Dichte 1 und alle anderen die Dichte 0 haben?

iii) Da sich in der Praxis oft ein überbestimmtes lineares Gleichungssystem mit einer sehr schlecht konditionierten Matrix ergibt, wendet man häufig die Tikhonov-Regularisierung an, um so das extreme Anwachsen von x, das aus Rauschkomponenten von v resultiert, zu verhindern. Aber auch auf das Gleichungssystem aus (i) kann man die Regularisierung anwenden. Wie sieht das regularisierte Gleichungssystem aus? Ist die Matrix des neuen Systems regulär?

iv) Nehmen Sie nun die Strahlen S_{10} bis S_{12} hinzu. Wie lässt sich die Dichteverteilung für den Messwert

$$v = (0, \sqrt{2}, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0)^T$$

bestimmen?

Zum Abschluss des Kapitels über lineare Gleichungssysteme ist nachfolgend wieder eine Aufgabe aus eine Semestralklausur aufgeführt. Diese ist zum Wiederholen des Stoffs gedacht und wird deshalb in der Übung nicht behandelt.

Wiederholung: Lineare Gleichungssysteme

a) Welche Aussagen sind wahr, welche falsch?

(1) Eine reelle Matrix Q ist genau dann orthogonal, wenn $Q^T = Q$ gilt.

(2) Gauß-Elimination mit Spaltenpivotsuche ist immer numerisch stabil.

(3) Das Verwenden von Pivotsuche bei der Berechnung der LR-Zerlegung kann die Kondition der Matrizen L und R deutlich verbessern.

(4) Das Lösen eines lineares Gleichungssystems der Gröše $n \times n$ mittels Gauš-Elimination benötigt im Allgemeinen ca. $\frac{2}{3}n^2$ Flops.

b) Gegeben seien nun die folgenden Messwerte (x_i, y_i):

i	1	2	3	4
x_i	-2	-1	0	1
y_i	2	2	3	7

Diese sollen mit einer Parabel

$$f(x) = a + bx + cx^2 \tag{9}$$

bestmglichst approximiert werden. Dazu sollen die Parameter a, b und c in (9) so bestimmt werden, dass der quadratische Fehler zwischen Messwerten und Parabelwerten, d.h.

$$\sum_{i=1}^{4} (f(x_i) - y_i)^2, \tag{10}$$

minimiert wird.

i) Stellen Sie zunächst das zugehörige überbestimmte lineare Gleichungssystem $Ax = b$ zur Bestimmung der Parameter a, b und c auf!

ii) Welches quadratische Gleichungssystem ist zu lösen, um die Koeffizienten a, b und c so zu bestimmen, dass der Fehler (10) minimiert wird?