

Lineare Gleichungssysteme und Ausgleichsprobleme sind Identifikationsprobleme für endlich viele Parameter. Sie treten auch im Zusammenhang mit linearen Operatorgleichungen in Funktionenräumen auf, wenn diese diskretisiert werden. Im Rahmen einer detaillierten Sensitivitätsanalyse wird die „Kondition“ von Ausgleichsproblemen als ein die Stabilitätsbedingung aus Definition 1.13 ersetzendes, feineres und für Anwendungen besser geeignetes Maß für Schlechtgestelltheit eingeführt. Die folgende Darstellung beschränkt sich auf $\mathbb{K} = \mathbb{R}$, eine Erweiterung auf $\mathbb{K} = \mathbb{C}$ wäre jedoch möglich.

2.1 Mathematischer Hintergrund

Ein lineares Gleichungssystem

$$Ax = b \quad \text{mit} \quad A \in \mathbb{R}^{m,n}, x \in \mathbb{R}^n, b \in \mathbb{R}^m, m \geq n, \quad (2.1)$$

ist ein Modell für ein inverses Problem: b steht für die „Wirkung“, x für die gesuchte „Ursache“ und der Kausalzusammenhang zwischen beiden wird durch die Abbildung $T : \mathbb{R}^n \rightarrow \mathbb{R}^m, x \mapsto Ax$, beschrieben. Wenn $m > n$, dann nennt man das Gleichungssystem **überbestimmt**. Ungenauigkeiten in den Komponenten von A oder b führen in der Regel zu Widersprüchen in den einzelnen Gleichungen eines überbestimmten Systems, so dass kein x existiert, welches (2.1) exakt erfüllt: Das **Residuum**

$$r(x) := b - Ax$$

verschwindet für kein $x \in \mathbb{R}^n$. Ersatzweise kann man nach einem x suchen, welches das Residuum wenigstens so klein wie möglich macht, etwa im Sinn der Euklidischen Norm:

$$\text{Finde } \hat{x} \text{ so, dass } \|r(\hat{x})\|_2 \leq \|r(x)\|_2 \quad \text{für alle } x \in \mathbb{R}^n. \quad (2.2)$$

Man nennt (2.2) das **lineare Ausgleichsproblem** oder auch die Auflösung der Widersprüche nach der **Methode der kleinsten Quadrate**. Äquivalent zu (2.2) ist die Minimierung von

$$f(x) = \|r(x)\|_2^2 = r(x)^T r(x) = x^T A^T A x - 2x^T A^T b + b^T b.$$

Die Ableitung nach x (der Gradient) lautet $\nabla f(x) = 2A^T A x - 2A^T b$. Eine notwendige Bedingung für einen Minimierer \hat{x} von (2.2) ist $\nabla f(\hat{x}) = 0$, also die Erfüllung der sogenannten **Normalengleichungen**

$$A^T A \hat{x} = A^T b \iff A^T \hat{r} = 0 \quad \text{mit} \quad \hat{r} := r(\hat{x}). \quad (2.3)$$

Diese Bedingungen sind auch hinreichend für eine Lösung. Wenn nämlich $x \in \mathbb{R}^n$ beliebig gewählt wird, dann ist $r(x) = \hat{r} + A(\hat{x} - x)$, also

$$\|r(x)\|_2^2 = \hat{r}^T \hat{r} + \underbrace{2(\hat{x} - x)^T A^T \hat{r}}_{=0} + (\hat{x} - x)^T A^T A (\hat{x} - x) \geq \|\hat{r}\|_2^2,$$

wobei $\|A(\hat{x} - x)\|_2 = 0 \Leftrightarrow A(\hat{x} - x) = 0 \Leftrightarrow r(x) = \hat{r}$. Dies beweist den

Satz 2.1 (Existenz und Eindeutigkeit der Lösung des Ausgleichsproblems) *Notwendig und hinreichend für einen Minimierer von (2.2) ist die Erfüllung der Normalengleichungen (2.3). Ein Minimierer \hat{x} ist genau dann eindeutig, wenn alle Spalten von A linear unabhängig sind, wenn also $\text{Rang}(A) = n$. Das Residuum \hat{r} ist immer eindeutig.*

Die Normalengleichungen haben eine geometrische Interpretation. Es ist

$$A = \left(\begin{array}{c|c|c|c} & & & \\ \hline & a_1 & a_2 & \cdots & a_n \\ \hline \end{array} \right), \quad \text{alle } a_j \in \mathbb{R}^m \implies A^T \hat{r} = \begin{pmatrix} a_1^T \hat{r} \\ \vdots \\ a_n^T \hat{r} \end{pmatrix},$$

somit bedeutet (2.3), dass das Residuum \hat{r} senkrecht auf $\mathcal{R}_A = \text{span}\{a_1, \dots, a_n\} = \{Ax; x \in \mathbb{R}^n\}$ steht, in Zeichen $\hat{r} \perp \mathcal{R}_A$, siehe Abb. 2.1.

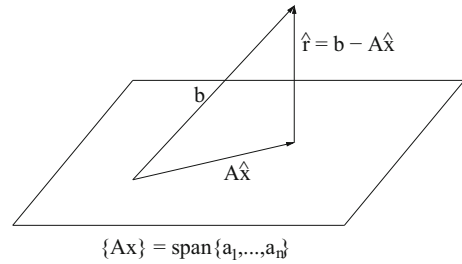
Beispiel 2.2 (Ausgleichsgerade) Es bestehe der Kausalzusammenhang

$$T : \mathbb{R} \rightarrow \mathbb{R}, \quad t \mapsto T(t) = \alpha + \beta(t - \gamma),$$

mit frei wählbarem Parameter γ und unbekannten, von γ abhängigen Parametern $\alpha, \beta \in \mathbb{R}$. Folgende Messwerte liegen vor (entnommen aus [5], Example 5.7.3):

t	1	3	4	6	7
$T(t)$	-2,1	-0,9	-0,6	0,6	0,9

Abb. 2.1 Lösung eines linearen Ausgleichsproblems



Wählt man $\gamma = 4$, dann ergibt sich folgendes überbestimmte lineare Gleichungssystem für $x = (\alpha, \beta)^T$:

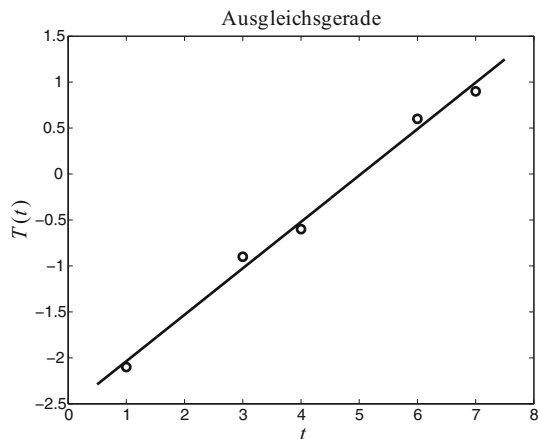
$$\underbrace{\begin{pmatrix} 1 & -3 \\ 1 & -1 \\ 1 & 0 \\ 1 & 2 \\ 1 & 3 \end{pmatrix}}_{=: A} \cdot \underbrace{\begin{pmatrix} \alpha \\ \beta \end{pmatrix}}_{=: b} = \underbrace{\begin{pmatrix} -2,1 \\ -0,9 \\ -0,6 \\ 0,6 \\ 0,9 \end{pmatrix}}_{=: b}.$$

Dieses Gleichungssystem ist nicht lösbar. Obwohl der Kausalzusammenhang zwischen t und $T(t)$ tatsächlich besteht, weichen die in der Tabelle angeführten Messwerte von den wahren Werten $T(t)$ ab, so dass widersprüchliche Gleichungen resultieren. Die Normalgleichungen lauten hier

$$\begin{pmatrix} 5 & 1 \\ 1 & 23 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} -2,1 \\ 11,1 \end{pmatrix}$$

mit der Lösung $\alpha \approx -0,52105$ und $\beta \approx 0,50526$. Abbildung 2.2 zeigt die so berechnete Ausgleichsgerade und dazu die markierten Messpunkte. \diamond

Abb. 2.2 Ausgleichsgerade zu Messwerten



Grundsätzlich könnte man das Minimierungsproblem (2.2) auch für andere Normen als für $\|\cdot\|_2$ betrachten und zum Beispiel $\|r(x)\|_1 = |r_1(x)| + \dots + |r_n(x)|$ oder $\|r(x)\|_\infty = \max\{|r_1(x)|, \dots, |r_n(x)|\}$ minimieren. Dies ist jedoch deutlich aufwändiger. Wenn die Inkonsistenzen in den Gleichungen von (2.1) zufällig sind, dann würde es außerdem dazu führen, dass maximale Inkonsistenzen („Ausreißer“) die Lösung stark beeinflussen, was meist unerwünscht ist. Eine mathematische Untersuchung von (2.1) bei zufälligen Inkonsistenzen findet man beispielsweise in Kapitel 17 von [23].

2.2 Sensitivitätsanalyse linearer Ausgleichsprobleme

Unter der Bedingung, dass die Matrix $A \in \mathbb{R}^{m,n}$ vollen Rang $n \leq m$ hat, besitzt das Ausgleichsproblem (2.2) eine eindeutige Lösung $\hat{x} = (A^T A)^{-1} A^T b$. Da lineare Abbildungen zwischen endlichdimensionalen Räumen immer stetig sind, handelt es sich dann bei (2.2) nach Definition 1.13 um ein wohlgestelltes Problem. In diesem Abschnitt wird genauer untersucht, wie sensitiv die Lösung von (2.2) von A und b abhängt. Als technisches Hilfsmittel der Analyse wird die Singulärwertzerlegung von A benutzt, siehe (A.1) im Anhang. Die (Eingabe-) Daten des linearen Ausgleichsproblems sind die Matrix A und die rechte Seite b . Das zu diesen Daten gehörige Resultat \hat{x} ist durch

$$\|b - A\hat{x}\|_2 \leq \|b - Ax\|_2 \quad \text{für alle } x \in \mathbb{R}^n \quad (2.4)$$

charakterisiert. Ändert man die Daten des Ausgleichsproblems hin zu einer Matrix $A + \delta A$ und einer rechten Seite $b + \delta b$, dann gehört dazu ein geändertes Resultat \tilde{x} , welches durch

$$\|(b + \delta b) - (A + \delta A)\tilde{x}\|_2 \leq \|(b + \delta b) - (A + \delta A)x\|_2 \quad \text{für alle } x \in \mathbb{R}^n \quad (2.5)$$

charakterisiert ist. In den folgenden beiden Sätzen werden Aussagen über den absoluten und den relativen Unterschied von \hat{x} und \tilde{x} in Abhängigkeit von den absoluten beziehungsweise relativen Änderungen der Eingabedaten gemacht.

Satz 2.3 (Absolute Sensitivität des Ausgleichsproblems) *Es seien $A, \delta A \in \mathbb{R}^{m,n}$ und $b, \delta b \in \mathbb{R}^m$, $m \geq n$. A habe singuläre Werte $\sigma_1 \geq \dots \geq \sigma_n > 0$, also vollen Rang. Es sei \hat{x} die Lösung von (2.2), erfülle also (2.4), sein Residuum sei $\hat{r} := b - A\hat{x}$. Unter der Voraussetzung*

$$\eta := \frac{\|\delta A\|_2}{\sigma_n} < 1, \quad \text{das heißt} \quad \|\delta A\|_2 < \sigma_n,$$

gibt es ein eindeutig bestimmtes $\tilde{x} \in \mathbb{R}^n$, welches (2.5) erfüllt und für $\delta x := \tilde{x} - \hat{x}$ gilt

$$\|\delta x\|_2 \leq \frac{1}{\sigma_n(1 - \eta)} \cdot (\|\delta b\|_2 + \|\delta A\|_2 \|\hat{x}\|_2) + \frac{1}{\sigma_n^2(1 - \eta)^2} \cdot \|\delta A\|_2 \cdot \|\hat{r}\|_2. \quad (2.6)$$

Beweis Die Matrix $A + \delta A$ hat vollen Rang, denn

$$\begin{aligned} (A + \delta A)x = 0 &\Rightarrow Ax = -\delta Ax \Rightarrow \sigma_n \|x\|_2 \leq \|Ax\|_2 \leq \|\delta A\|_2 \|x\|_2 \\ &\Rightarrow \underbrace{(\sigma_n - \|\delta A\|_2)}_{> 0} \|x\|_2 \leq 0, \end{aligned}$$

was nur für $x = 0$ möglich ist. Nach Satz 2.1 ist \tilde{x} eindeutig bestimmt.

Mit $\tilde{A} = A + \delta A$ und $\tilde{b} = b + \delta b$ gilt nach Satz 2.1 und mit $M := (\tilde{A}^T \tilde{A})^{-1}$, dass $\tilde{x} = M \tilde{A}^T \tilde{b}$ (Normalengleichungen). Folglich ist

$$\begin{aligned} \delta x &= \tilde{x} - \hat{x} = M \tilde{A}^T \tilde{b} - \hat{x} = M \tilde{A}^T (\tilde{b} - \tilde{A} \hat{x}) \\ &= M \tilde{A}^T (b - A \hat{x}) + M \tilde{A}^T (\delta b - \delta A \hat{x}) \\ &= M(\delta A)^T \hat{r} + M \tilde{A}^T (\delta b - \delta A \hat{x}) \quad [(A + \delta A)^T \hat{r} = (\delta A)^T \hat{r} \text{ nach (2.3)}] \end{aligned}$$

und daraus ergibt sich

$$\|\delta x\|_2 \leq \|M\|_2 \|\delta A\|_2 \|\hat{r}\|_2 + \|M \tilde{A}^T\|_2 (\|\delta b\|_2 + \|\delta A\|_2 \|\hat{x}\|_2).$$

Alles ist bewiesen, wenn $\|M\|_2 \leq 1/(\sigma_n^2(1 - \eta)^2)$ und $\|M \tilde{A}^T\|_2 \leq 1/(\sigma_n(1 - \eta))$. Aus einer SVD $\tilde{A} = \tilde{U} \tilde{\Sigma} \tilde{V}^T$ erhält man

$$M = \tilde{V} (\tilde{\Sigma}^T \tilde{\Sigma})^{-1} \tilde{V}^T = \tilde{V} \text{diag}(1/\tilde{\sigma}_1^2, \dots, 1/\tilde{\sigma}_n^2) \tilde{V}^T$$

und daraus $M \tilde{A}^T = \tilde{V} \tilde{\Sigma}^+ \tilde{U}^T$, wobei $\tilde{\Sigma}^+ = \text{diag}(1/\tilde{\sigma}_1, \dots, 1/\tilde{\sigma}_n) \in \mathbb{R}^{n,m}$. Hieraus folgt $\|M\|_2 = 1/\tilde{\sigma}_n^2$ sowie $\|M \tilde{A}^T\|_2 = 1/\tilde{\sigma}_n$. Nach Satz A.1 ist $|\sigma_n - \tilde{\sigma}_n| \leq \|\delta A\|_2$, also ist $\tilde{\sigma}_n \geq \sigma_n - \|\delta A\|_2 = \sigma_n(1 - \eta)$ und somit $\|M \tilde{A}^T\|_2 \leq 1/(\sigma_n(1 - \eta))$. Analog für $\|M\|_2$. \square

Die im Satz angegebene Schranke für $\|\delta x\|_2$ ist „fast scharf“: in [2], S. 29, wird ein Beispiel angegeben, bei dem sie näherungsweise erreicht wird.

Satz 2.4 (Relative Sensitivität des Ausgleichsproblems) *Es seien $A, \delta A \in \mathbb{R}^{m,n}$ und $b, \delta b \in \mathbb{R}^m$, $m \geq n$. A habe singuläre Werte $\sigma_1 \geq \dots \geq \sigma_n > 0$ und es sei*

$$\kappa_2(A) := \frac{\sigma_1}{\sigma_n}. \quad (2.7)$$

Es sei $\hat{x} \neq 0$ die Lösung von (2.2), erfülle also (2.4), sein Residuum sei $\hat{r} := b - A\hat{x}$. Ferner gelte für ein $\varepsilon > 0$

$$\|\delta A\|_2 \leq \varepsilon \|A\|_2, \quad \|\delta b\|_2 \leq \varepsilon \|b\|_2 \quad \text{und} \quad \kappa_2(A)\varepsilon < 1. \quad (2.8)$$

Dann gibt es ein durch (2.5) eindeutig bestimmtes $\tilde{x} \in \mathbb{R}^n$ und für $\delta x := \tilde{x} - \hat{x}$ gilt

$$\frac{\|\delta x\|_2}{\|\hat{x}\|_2} \leq \frac{\kappa_2(A)\varepsilon}{1 - \kappa_2(A)\varepsilon} \left(2 + \left(\frac{\kappa_2(A)}{1 - \kappa_2(A)\varepsilon} + 1 \right) \frac{\|\hat{r}\|_2}{\|A\|_2 \|\hat{x}\|_2} \right). \quad (2.9)$$

Beweis Aus der SVD von A leitet sich die Identität $\sigma_1 = \|A\|_2$ ab. Also folgt aus $\|\delta A\|_2 \leq \varepsilon \|A\|_2$, dass $\eta = \|\delta A\|_2 / \sigma_n \leq \kappa_2(A) \varepsilon$. Unter der Voraussetzung $\kappa_2(A) \varepsilon < 1$ ist demnach die Bedingung $\eta < 1$ aus Satz 2.3 erst recht erfüllt und die Abschätzung (2.6) bleibt auch gültig, wenn man η durch $\kappa_2(A) \varepsilon$ ersetzt. Division durch $\|\hat{x}\|_2$ liefert

$$\frac{\|\delta x\|_2}{\|\hat{x}\|_2} \leq \frac{1}{\sigma_n(1 - \kappa_2(A) \varepsilon)} \left(\frac{\|\delta b\|_2}{\|\hat{x}\|_2} + \|\delta A\|_2 \right) + \frac{1}{(\sigma_n(1 - \kappa_2(A) \varepsilon))^2} \cdot \frac{\|\delta A\|_2 \|\hat{r}\|_2}{\|\hat{x}\|_2}.$$

Wegen $\|\delta A\|_2 \leq \varepsilon \|A\|_2$, $\|\delta b\|_2 \leq \varepsilon \|b\|_2$ und $\|A\|_2 / \sigma_n = \kappa_2(A)$ ergibt sich

$$\frac{\|\delta x\|_2}{\|\hat{x}\|_2} \leq \frac{\kappa_2(A) \varepsilon}{1 - \kappa_2(A) \varepsilon} \left(\frac{\|b\|_2}{\|A\|_2 \|\hat{x}\|_2} + 1 \right) + \frac{\kappa_2(A)^2 \varepsilon}{(1 - \kappa_2(A) \varepsilon)^2} \cdot \frac{\|\hat{r}\|_2}{\|A\|_2 \|\hat{x}\|_2}.$$

Gleichung (2.9) folgt aus $\|b\|_2 \leq \|b - A\hat{x}\|_2 + \|A\hat{x}\|_2 \leq \|\hat{r}\|_2 + \|A\|_2 \|\hat{x}\|_2$. \square

In den Sätzen 2.3 und 2.4 wird die Sensitivität des Ausgleichsproblems *quantifiziert* und ein Maß dafür angegeben, wie empfindlich dessen Resultat auf Änderungen in den Daten A und b reagiert. Insbesondere gilt $\delta x \rightarrow 0$ für $\delta A \rightarrow 0$ und $\delta b \rightarrow 0$. Erneut lässt sich also feststellen, dass das lineare Ausgleichsproblem wohlgestellt ist nach Definition 1.13. In beiden voranstehenden Sätzen ist eine Aussage über lineare Gleichungssysteme enthalten, da die Lösung \hat{x} des Systems (2.1), wenn sie eindeutig existiert, identisch ist mit der von (2.2) – es ist dann $\hat{r} = 0$. Entscheidend für die Sensitivität ist erstens die Größe $\kappa_2(A) \varepsilon$, die kleiner als 1 sein muss. Bei genügend kleinem ε ist $\kappa_2(A) / (1 - \kappa_2(A) \varepsilon) \approx \kappa_2(A)$, so dass (näherungsweise) $\kappa_2(A)$ als Verstärkungsfaktor von relativen Fehlern in den Eingabedaten auftritt. Bei (eindeutig lösbar) linearen Gleichungssystemen kommt es nur auf diesen Faktor an, weil dort $\hat{r} = 0$ gilt. Bei linearen Ausgleichsproblemen kommt es zweitens hingegen auch auf \hat{r} und damit auf die rechte Seite b an, denn es tritt $\kappa_2(A)^2 \|\hat{r}\|_2$ ebenfalls als Verstärkungsfaktor relativer Datenfehler auf.

Das inverse Problem (2.1) beziehungsweise (2.2) wird als **gut konditioniert** (**schlecht konditioniert**) bezeichnet, wenn kleine Änderungen in den Daten A und b nur kleine (auch große) Änderungen in der Lösung bewirken können. Die **Konditionszahl** ist der Faktor, um den sich relative Änderungen in den Daten A oder b schlimmstenfalls in den entsprechenden relativen Änderungen des Resultats verstärken können. Für die Kondition eines linearen Gleichungssystems (2.1) ist (näherungsweise) die Zahl $\kappa_2(A)$ bestimmend, beim linearen Ausgleichsproblems (2.2) ist es das Maximum der Zahlen $\kappa_2(A)$ und $\kappa_2(A)^2 \|\hat{r}\|_2 / (\|A\|_2 \|\hat{x}\|_2)$. Die Zahl $\kappa_2(A) = \sigma_1 / \sigma_n$ heißt **Konditionszahl der Matrix A** .

Wenn die Matrix A *keinen* vollen Rang mehr hat, wenn also $\sigma_n = 0$ gilt, könnte man formal

$$\kappa_2(A) = \sigma_1 / 0 = \infty$$

setzen und von „unendlich schlechter Kondition“ des Ausgleichsproblems sprechen im Einklang damit, dass dann keine eindeutige Lösung von (2.2) mehr existiert. Die Eindeutigkeit einer Lösung lässt sich weiterhin erzwingen, wenn man die *zusätzliche Bedingung* stellt, dass unter allen Lösungen von (2.2) diejenige mit der kleinsten Norm ausgewählt werden soll. Dies ist eine gegenüber (2.2) *geänderte Problemstellung*, die sich mathematisch folgendermaßen ausdrücken lässt:

$$\text{Finde } \hat{x} \text{ so, dass } \|\hat{x}\|_2 \leq \|x\|_2 \text{ für alle } x \in M := \arg \min\{\|b - Ax\|_2\}, \quad (2.10)$$

wobei $\arg \min\{\|b - Ax\|_2\}$ die Menge aller Lösungen von (2.2) ist. Man nennt eine Lösung von (2.10) **Minimum-Norm-Lösung** des Ausgleichsproblems (2.2). Im Fall $\text{Rang}(A) = n$ hat (2.2) eine eindeutige Lösung \hat{x} und diese ist auch die Lösung von (2.10). Der folgende Satz zeigt zum einen, dass für jede beliebige Matrix A eine eindeutige Lösung von (2.10) existiert. Er zeigt zum anderen, dass das Problem (2.10) besser konditioniert sein kann als (2.2).

Satz 2.5 (Ausgleichsproblem bei singulärer Matrix) $A \in \mathbb{R}^{m,n}$ habe den Rang $r \leq n \leq m$ und seine SVD laute

$$\begin{aligned} A &= U \Sigma V^T = \left(\underbrace{U_1}_r, \underbrace{U_2}_{n-r}, \underbrace{U_3}_{m-n} \right) \underbrace{\begin{pmatrix} \Sigma_1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}}_{\substack{r \quad n-r}} \underbrace{\left(\underbrace{V_1}_r, \underbrace{V_2}_{n-r} \right)^T}_{m-n} \\ &= U_1 \Sigma_1 V_1^T. \end{aligned}$$

(a) Alle Lösungen von (2.2) haben die Form

$$x = V_1 \Sigma_1^{-1} U_1^T b + V_2 z, \quad z \in \mathbb{R}^{n-r} \text{ beliebig.} \quad (2.11)$$

(b) Unter allen Lösungen gibt es genau eine mit minimaler Euklidischer Norm, also eine eindeutige Lösung von (2.10). Diese erhält man für $z = 0$, also

$$x = V_1 \Sigma_1^{-1} U_1^T b.$$

Für dieses x ist $\|x\|_2 \leq \|b\|_2 / \sigma_r$.

(c) Ändert man b zu $b + \delta b$, so erhält man dazu eine eindeutige Lösung $x + \delta x$ von (2.10). Für diese ist

$$\|\delta x\|_2 \leq \frac{\|\delta b\|_2}{\sigma_r}.$$

Beweis Teil (a):

$$\begin{aligned}\|b - Ax\|_2^2 &= \left\| U^T b - \begin{pmatrix} U_1^T \\ U_2^T \\ U_3^T \end{pmatrix} U_1 \Sigma_1 V_1^T x \right\|_2^2 = \left\| \begin{pmatrix} U_1^T b - \Sigma_1 V_1^T x \\ U_2^T b \\ U_3^T b \end{pmatrix} \right\|_2^2 \\ &= \|U_1^T b - \Sigma_1 V_1^T x\|_2^2 + \|U_2^T b\|_2^2 + \|U_3^T b\|_2^2\end{aligned}$$

wird minimiert, wenn

$$\Sigma_1 V_1^T x = U_1^T b \iff x = V_1 \Sigma_1^{-1} U_1^T b + V_2 z$$

für beliebiges $z \in \mathbb{R}^{n-r}$, denn

$$\mathcal{N}_{V_1^T} = \mathcal{R}_{V_1}^\perp = \{V_2 z; z \in \mathbb{R}^{n-r}\}.$$

Teil (b): Da die Spalten von V_1 und V_2 paarweise senkrecht aufeinander stehen, folgt aus (2.11) mit dem Satz des Pythagoras, dass

$$\|x\|_2^2 = \|V_1 \Sigma_1^{-1} U_1^T b\|_2^2 + \|V_2 z\|_2^2$$

und dieser Ausdruck wird genau dann minimal, wenn $V_2 z = 0$, also genau dann, wenn $z = 0$. Für das entsprechende x bekommt man

$$\|x\|_2^2 = \|V_1 \Sigma_1^{-1} U_1^T b\|_2^2 = \left\| \begin{pmatrix} u_1^T b / \sigma_1 \\ \vdots \\ u_r^T b / \sigma_r \end{pmatrix} \right\|_2^2 \leq \frac{1}{\sigma_r^2} \sum_{j=1}^r |u_j^T b|^2 \leq \frac{\|b\|_2^2}{\sigma_r^2}$$

Teil (c): Mit $b + \delta b$ statt b bekommt man nach Teil (b) die Minimum-Norm-Lösung $x + \delta x = V_1 \Sigma_1^{-1} U_1^T (b + \delta b)$, also $\delta x = V_1 \Sigma_1^{-1} U_1^T \delta b$ und damit ergibt sich die Behauptung wie in (b). \square

Die Empfindlichkeit der Minimum-Norm-Lösung gegenüber Ungenauigkeiten im Vektor b ist durch die Größe des kleinsten nicht verschwindenden singulären Wertes σ_r von A bestimmt. Werden auch Störungen in A berücksichtigt, dann kann man Abschätzungen für $\|\delta x\|_2$ wie in Satz 2.3 zeigen mit σ_r statt σ_n , siehe etwa Theorem 1.4.6 in [2].

Definition 2.6 (Pseudoinverse) *Unter den Voraussetzungen des Satzes 2.5 heißt*

$$A^+ := V_1 \Sigma_1^{-1} U_1^T$$

Pseudoinverse von A .

Die stets eindeutige Minimum-Norm-Lösung des linearen Ausgleichsproblems lässt sich nach Satz 2.5 in der Form

$$x = A^+ b$$

schreiben. Im Fall $\text{Rang}(A) = n$ gilt $A^+ = (A^T A)^{-1} A^T$ in Übereinstimmung mit der dann eindeutigen Lösung der Normalgleichungen. Wenn $\text{Rang}(A) = n$ und zusätzlich $m = n$, dann ist $A^+ = A^{-1}$ in Übereinstimmung mit der dann eindeutigen Lösung des linearen Gleichungssystems $Ax = b$.

Beispiel 2.7 Für $\varepsilon > 0$ seien

$$A = \begin{pmatrix} 1 & 0 \\ 0 & \varepsilon \end{pmatrix} \text{ und } b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \implies A^+ = \begin{pmatrix} 1 & 0 \\ 0 & \varepsilon^{-1} \end{pmatrix}, A^+ b = \begin{pmatrix} b_1 \\ b_2/\varepsilon \end{pmatrix}.$$

Bei der Lösung des Ausgleichsproblems (2.2) beziehungsweise des Minimum-Norm-Problems (2.10), welche beide in diesem Fall nur verkappte lineare Gleichungssysteme sind, werden Fehler in der Komponente b_2 der rechten Seite um einen Faktor ε^{-1} verstärkt. Da $\kappa_2(A) = \varepsilon^{-1}$, deckt sich diese Aussage mit der von Satz 2.4. Die Kondition der Lösung von (2.2) wird mit $\varepsilon \rightarrow 0$ beliebig schlecht. Sei nun andererseits

$$\begin{aligned} B &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}_{=U} \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}}_{=\Sigma} \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}^T}_{=V^T} \\ &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} 1 \begin{pmatrix} 1 \\ 0 \end{pmatrix}^T \implies B^+ = B. \end{aligned}$$

Die Matrix B hat Rang 1 und $\sigma_1 = 1$ ist ihr kleinster nicht verschwindender Eigenwert. Die Lösung $B^+ b = (b_1, 0)^T$ des Minimum-Norm-Problems (2.10) ist in Bezug auf Änderungen der rechten Seite hervorragend konditioniert. \diamond

Inverse Probleme

Grundlagen, Theorie und Anwendungsbeispiele

Richter, M.

2015, IX, 128 S. 24 Abb., Softcover

ISBN: 978-3-662-45810-5