R. Hiptmair
S. Pintarelli
E. Spindler

Herbstsemester 2014

ETH Zürich
D-MATH

Lineare Algebra und Numerische Mathematik für D-BAUG

Beispiellösung für Serie 11

Einleitung. Der Schwerpunkt dieses Übungsblattes ist auf numerischer linearer Algebra in MATLAB. Daneben werden auch wieder die Themen „Lineare Ausgleichsrechnung" und „Orthogonalität" wieder aufgegriffen.

(★★) Aufgabe 11.1 Normalengleichungen ohne Matrixprodukt

In [LANM, Abschnitt 4.4] wurde das Konzept der Kleinste-Quadrate-Lösung von überbestimmten linearen Gleichungssuystemen eingeführt und mit den Normalengeleichungen aus [LANM, Abschnitt 4.4.3] und den Orthogonalisierungstechniken aus [LANM, Abschnitt 4.4.4] wurden zwei praktische Verfahren zur Berechnung von Kleinste-Quadrate-Lösungen vorgestellt.

In dieser Aufgabe geht es nun um eine Umformulierung der *Normalengleichungen*, die für überbestimmte lineare Gleichungssysteme mit relativ vielen Unbekannten und *dünnbesetzten Koeffizientenmatrizen* wichtig ist.

Für n, $m \in \mathbb{N}$ ist gegeben eine Matrix $\mathbf{A} \in \mathbb{R}^{m,n}$ mit $m > n$ und $\mathrm{Rang}(\mathbf{A}) = n$ und ein Rechte-Seite-Vektor $\mathbf{b} \in \mathbb{R}^n$. Wir betrachten nun das folgende Gleichungssystem, das auch als *erweiterte Normalengleichung* zum überbestimmten linearen Gleichungssystem $\mathbf{Ax} = \mathbf{b}$ bezeichnet wird:

$$\begin{bmatrix} \mathbf{A}^\top & \mathbf{0} \\ \mathbf{I} & \mathbf{A} \end{bmatrix} \begin{bmatrix} \mathbf{r} \\ \mathbf{x} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{b} \end{bmatrix}. \tag{11.1.1}$$

(★) **11.1a)** Was ist die Grösse der Koeffizientenmatrix von (11.1.1).

(★) **11.1b)** In welcher Beziehung steht die Lösungskomponente $\mathbf{r}$ aus (11.1.1) zur Lösungskomponente $\mathbf{x}$?

(→ Tipp)

Lösung: Aus Teilaufgabe 11.1c) wissen wir bereits, dass gilt $\mathbf{r} = \mathbf{b} - \mathbf{A}$. Er repräsentiert somit das *Residuum* aus der Definition der Kleinste-Quadrate-Lösung [LANM, Thm. IV.4.2.B] der Vorlesung.

(★★) **11.1c)** Zeigen Sie, dass gilt: $\mathbf{x} \in \mathbb{R}^m$ löst die Normalengleichungen $\mathbf{A}^\top\mathbf{Ax} = \mathbf{A}^\top\mathbf{b}$ zum überbestimmten linearen Gleichungssystem $\mathbf{Ax} = \mathbf{b}$ *genau dann*, wenn es einen Vektor $\mathbf{r} \in \mathbb{R}^n$ so gibt, dass $\left[\begin{smallmatrix} \mathbf{r} \\ \mathbf{x} \end{smallmatrix}\right]$ die erweiterten Normalengleichungen (11.1.1) löst.

Tipp: Es sind zwei Richtungen zu zeigen:

1. Ist $\left[\begin{smallmatrix} \mathbf{r} \\ \mathbf{x} \end{smallmatrix}\right]$ Lösung von (11.1.1), dann löst $\mathbf{x}$ die Normalengleichungen $\mathbf{A}^\top\mathbf{Ax} = \mathbf{A}^\top\mathbf{b}$ zum überbestimmten linearen Gleichungssystem $\mathbf{Ax} = \mathbf{b}$.

2. Löst $\mathbf{x}$ die Normalengleichungen $\mathbf{A}^\top\mathbf{A}\mathbf{x} = \mathbf{A}^\top\mathbf{b}$ zum überbestimmten linearen Gleichungssystem $\mathbf{A}\mathbf{x} = \mathbf{b}$, dann kann man (mit Hilfe der zweiten Gleichung von (11.1.1)) einen Vektor $\mathbf{r}$ so definieren, dass $\left[\begin{smallmatrix}\mathbf{r}\\ \mathbf{x}\end{smallmatrix}\right]$ (11.1.1) löst.

(→ Tipp)

Lösung: Wir zeigen zuerst, dass folgende Aussage gilt:

Sofern gilt

$$\begin{bmatrix}\mathbf{A}^\top & \mathbf{0}\\ \mathbf{I} & \mathbf{A}\end{bmatrix}\begin{bmatrix}\mathbf{r}\\ \mathbf{x}\end{bmatrix} = \begin{bmatrix}\mathbf{0}\\ \mathbf{b}\end{bmatrix}, \tag{11.1.2}$$

dann folgt $\mathbf{A}^\top\mathbf{A}\mathbf{x} = \mathbf{A}^\top\mathbf{b}$. Was nun folgt ist eine Blockelimination, die hier in einfache Schritte zerlegt wird.

Multiplizieren wir (11.1.2) aus, so erhalten wir

$$\mathbf{A}^\top\mathbf{r} = \mathbf{0}$$
$$\mathbf{r} + \mathbf{A}\mathbf{x} = \mathbf{b}\,.$$

Multiplizieren wir $\mathbf{r} + \mathbf{A}\mathbf{x} = \mathbf{b}$ von links mit $\mathbf{A}^\top$, so erhalten wir weiter

$$\mathbf{A}^\top\mathbf{r} + \mathbf{A}^\top\mathbf{A}\mathbf{x} = \mathbf{A}^\top\mathbf{b}\,.$$

Wir setzen die Bedingung $\mathbf{A}^\top\mathbf{r} = \mathbf{0}$ in die vorangehende Gleichung ein und bekommen die gesuchte Gleichung

$$\mathbf{A}^\top\mathbf{A}\mathbf{x} = \mathbf{A}^\top\mathbf{b}\,.$$

Umgekehrt nehmen wir an, dass gilt

$$\mathbf{A}^\top\mathbf{A}\mathbf{x} = \mathbf{A}^\top\mathbf{b}\,.$$

Wir definieren $\mathbf{r} := \mathbf{A}\mathbf{x} - \mathbf{b}$. Erinnern Sie sich daran, dass dies tatsächlich erlaubt ist, da wir $\mathbf{A}, \mathbf{x}$ und $\mathbf{b}$ bereits kennen. Somit bleibt zu zeigen, dass $\mathbf{r}$ die Bedingung

$$\mathbf{A}^\top\mathbf{r} = \mathbf{0}$$

erfüllt ist.

Nach Annahme gilt bereits

$$\mathbf{A}^\top\mathbf{r} = \mathbf{A}^\top\mathbf{A}\mathbf{x} - \mathbf{A}^\top\mathbf{b} = \mathbf{0}\,,$$

da $\mathbf{x}$ per Definition die Lösung der Normalengleichungen darstellt.

Somit haben wir gezeigt, dass beide Implikationen/Richtungen gelten und wir sind fertig mit dem Beweis.

(★★) **11.1d)** Welchen Vorteil kann es haben, das Gleichungssystem (11.1.1) anstelle der eigentlichen Normalengleichungen $\mathbf{A}^\top\mathbf{A}\mathbf{x} = \mathbf{A}^\top\mathbf{b}$ zu lösen?

(→ Tipp)

Lösung: In Teilaufgabe 10.9d) und 10.9e) war die Matrix $\mathbf{A}$ des überbestimmten Systems "dünnbesetzt", wohingegen die Matrix $\mathbf{A}^\top\mathbf{A}$ aus den zugehörigen Normalengleichungen eine "dichtbesetzte" Matrix war. Deshalb ist es ein grosser Vorteil, wenn wir $\left[\begin{smallmatrix}\mathbf{A}^\top & \mathbf{0}\\ \mathbf{I} & \mathbf{A}\end{smallmatrix}\right]\left[\begin{smallmatrix}\mathbf{r}\\ \mathbf{x}\end{smallmatrix}\right] = \left[\begin{smallmatrix}\mathbf{0}\\ \mathbf{b}\end{smallmatrix}\right]$ anstelle der Normalengleichungen betrachten, da dort die Matrix $\left[\begin{smallmatrix}\mathbf{A}^\top & \mathbf{0}\\ \mathbf{I} & \mathbf{A}\end{smallmatrix}\right]$ "dünnbesetzt" bleibt, sofern die Matrix $\mathbf{A}$ selbst "dünnbesetzt" ist.

Aufgabe 11.2 Lineare Regression in der Signalverarbeitung

Die lineare Ausgleichsrechnung aus [LANM, Thm. 4.4] spielt auch eine grosse Rolle in der modernen Signalverarbeitung, zum Beispiel wenn es um das Ausmessen von Übertragungskanälen geht. Man erhält auf diese Weise etwa die Werte für die Gewichte in der Point-Spread-Function, die die Verzerrung von Pixelbildern in optischen Übertragungssystemen beschreibt, siehe [LANM, Abschnitt 5.6.2]. In dieser Aufgabe behandeln wir aber einen einfacheren, „eindimensionalen" Fall, nämlich die Störung eines zeitdiskreten Signals bei der Übertragung.

Wir betrachten ein zeitdiskretes Signal, welches gegeben ist durch die Folge von Werten $x_1, \ldots, x_m$, $m \in \mathbb{N}$, die die Signalamplituden zu festen Abtastzeitpunkten angeben. Diese Signale werden durch einen *linearen* Übertragungkanal (in Abbildung 11.1 mit einem blauen Pfeil symbolisiet) geschickt, wobei vorher und nachher "Funktstille" herrscht.

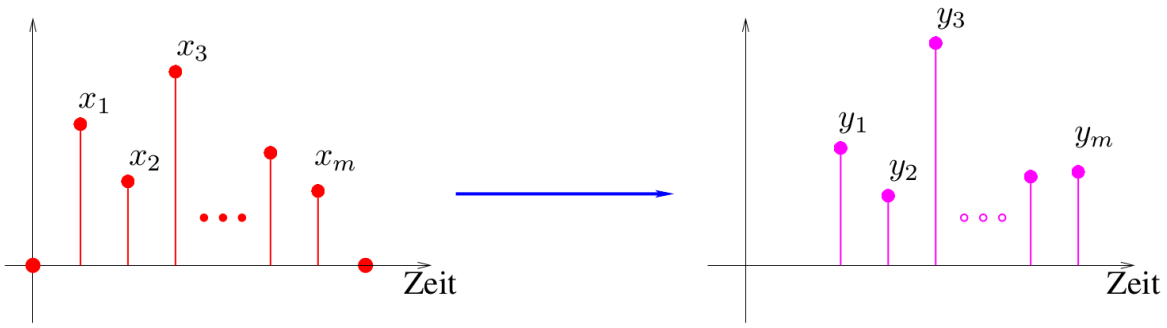

Abbildung 11.1: Übertragungskanal und zeitdiskrete Signale auf Senderseite (links) und Empfängerseite (rechts)

Leider ist der Kanal nicht perfekt, so dass sich zeitlich benachbarte Signalwerte gegenseitig beeinflussen, ein Phänomen, das man als Übersprechen bezeichnet. Daher besteht das am anderen Ende des Kanals empfangene Signal aus einer anderen Folge von Werten $y_1, \ldots, y_m$.

Der Zusammenhang zwischen dem empfangenen und gesendeten Signal ist näherungsweise gegeben durch

$$y_i = \beta x_{i-1} + \alpha x_i + \beta x_{i+1}\,, \quad i = 1, \ldots, m\,, \tag{11.2.1}$$

wobei wir annehmen, dass gilt $x_0 = x_{m+1} = 0$.

Die Aufgabe besteht nun darin, die unbekannten Übersprechparameter α und β aus den *gemessenen* Signalen zu ermitteln und zwar durch Lösen eines überbestimmten linearen Gleichungssystems mit der Methode der kleinsten Quadrate aus [LANM, Abschnitt 4.4.2] der Vorlesung.

⊛ **11.2a)** Stellen Sie das überbestimmte lineare Gleichungssystem für das Schätzproblem auf.

(→ Tipp)

Lösung: Setze

$$\mathbf{A} = \begin{bmatrix} x_1 & x_2 \\ x_2 & x_1 + x_3 \\ x_3 & x_2 + x_4 \\ \vdots & \vdots \\ x_{m-1} & x_{m-2} + x_m \\ x_m & x_{m-1} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \quad \text{und} \quad \mathbf{b} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}.$$

Dann ist unser überbestimmtes Gleichungssystem $\mathbf{Ax} = \mathbf{b}$.

(★★) **11.2b)** Schreiben Sie eine MATLAB-Funktion

```
function [beta, alpha] = CrosstalkChannel(x, y),
```

welche für die Eingabedaten $x_1, \ldots, x_m$ und $y_1, \ldots, y_m$, repräsentiert durch die Spaltenvektoren `x` und `y`, die Werte für α und β aus (11.2.1) im Sinne der kleinsten Quadrate bestimmt.

(→ Tipp)

Lösung:

```
?% Problem 6c)
% INPUT
% x, y - signals, of equal length
% OUTPUT
% beta, alpha - parameters which give an appropriate least
   squares solution

function [beta, alpha] = CrosstalkChannel(x, y)

% Here we are addapting the linearregression.m code
n = length(y); if (length(x) ~= n), error('data size
   mismatch'); end

%B Build the matrix of the overdetermined system
A = [x, [x(2); x(1:n-2)+x(3:n);x(n-1)] ];
% Compute the solution
solution = A\y;
% Assign appropriate values
alpha = solution(1);
beta = solution(2);

end
```

11.2c) Geben Sie die Normalengleichungen gemäss [LANM, Satz IV.4.3.B] für das überbestimmte lineare Gleichungssystem aus Teilaufgabe 11.2a) an.

Lösung: Wir schreiben

$$\mathbf{C} = \mathbf{A}^\top \mathbf{A} = \begin{bmatrix} x_1 & x_2 & \dots & x_m \\ x_2 & x_1 + x_3 & \dots & x_{m-1} \end{bmatrix} \begin{bmatrix} x_1 & x_2 \\ x_2 & x_1 + x_3 \\ x_3 & x_2 + x_4 \\ \vdots & \vdots \\ x_m & x_{m-1} \end{bmatrix}.$$

Dann gilt $c_{11} = \sum_{i=1}^{m} x_i^2$. Weiter gilt

$$c_{12} = c_{21} = x_1 x_2 + \sum_{i=2}^{m-1} x_i(x_{i-1} + x_{i+1}) + x_{m-1}x_m = x_1 x_2 + \sum_{i=2}^{m-1} x_i x_{i-1} + \sum_{i=2}^{m-1} x_i x_{i+1} + x_m x_{m-1}$$
$$= \sum_{i=2}^{m} x_i x_{i-1} + \sum_{i=1}^{m-1} x_i x_{i+1} = 2 \sum_{i=1}^{m-1} x_i x_{i+1}.$$

Für den letzten Eintrag in $\mathbf{C}$ erhalten wir

$$c_{22} = x_2^2 + \sum_{i=2}^{m-1} (x_{i-1} + x_{i+1})^2 + x_m^2 = 2 \sum_{i=2}^{m-2} x_i^2 + 2 \sum_{i=2}^{m-1} x_{i-1} x_{i+1} + x_1^2 + x_m^2.$$

Berechnen wir nun $\mathbf{A}^\top \mathbf{b}$. Wir erhalten

$$\mathbf{A}^\top \mathbf{b} = \begin{bmatrix} \sum_{i=1}^{m} x_i y_i \\ x_2 y_1 + x_{m-1} y_m + \sum_{i=2}^{m-1} y_i (x_{i-1} + x_{i+1}) \end{bmatrix}$$

(★★★) **11.2d)** Für welche Signale $x_1, x_2, \dots, x_m$ können wir keine eindeutige Kleinste-Quadrate-Lösung erwarten? Hier geben theoretische Resultate aus der linearen Algebra dem Ingenieur Hinweise darauf, wie der die Messung durchzuführen hat.

Tipp: Die Lösung dieser anspruchsvollen Teilaufgabe erfordert Techniken aus [LANM, Abschnitt 7.4]!

(→ Tipp)

Lösung: Nach [LANM, Kor. IV.4.3.F] muss $\mathrm{Rang}(\mathbf{A}) = 2$ sein, damit die Normalengleichungen eine eindeutige Lösung besitzen. Wir bringen $\beta x_{i-1} + \alpha x_i + \beta x_{i+1}$ in Matrixform:

$$\mathbf{M}\mathbf{x} := \begin{bmatrix} \alpha & \beta & & & 0 \\ \beta & \alpha & \beta & & \\ & \beta & \alpha & \ddots & \\ & & \ddots & \ddots & \beta \\ 0 & & & \beta & \alpha \end{bmatrix} \mathbf{x} = \mathbf{y}$$

Falls nun $\mathbf{x}^\ell$ ein Eigenvektor von $\mathbf{M}$ ist, gilt $\mathbf{M}\mathbf{x}^\ell = \lambda_\ell \mathbf{x}^\ell = \mathbf{y}$. Wir folgen dem Tipp und machen den Ansatz:

$$\mathbf{x}^\ell = (x_j^\ell)_{j=1}^m \in \mathbb{R}^m \,, \quad \ell \in \{1, \dots, m\} \quad \text{mit} \quad x_j^\ell := \sin(\pi \frac{j\ell}{m+1}) \,, \quad j = 1, \dots, m \,. \quad (11.2.2)$$

Einsetzen in (11.2.1) ergibt:

$$\beta \mathbf{x}_{i-1}^{\ell} + \alpha \mathbf{x}_{i}^{\ell} + \beta \mathbf{x}_{i+1}^{\ell} = y_i \tag{11.2.3}$$

$$\Leftrightarrow \ \beta \sin\left(\pi \frac{(i-1)\ell}{m+1}\right) + \alpha \sin\left(\pi \frac{i\ell}{m+1}\right) + \beta \sin\left(\pi \frac{(i+1)\ell}{m+1}\right) = \mathbf{y}_i \tag{11.2.4}$$

Als nächstes verwendet man die trigonometrischen Identitäten $\sin(a+b) = \sin(a)\cos(b) + \cos(a)\sin(b)$ und $\sin(a-b) = \sin(a)\cos(b) - \cos(a)\sin(b)$. Damit wird aus (11.2.4) für $a = \pi\frac{i\ell}{m+1}$ und $b = \pm\pi\frac{\ell}{m+1}$:

$$\begin{aligned} 2\beta \sin\left(\frac{\pi i\ell}{m+1}\right)\cos\left(\frac{\pi\ell}{m+1}\right) + \alpha sin\left(\frac{\pi i\ell}{m+1}\right) &= 2\beta \mathbf{x}_i^{\ell}\cos\left(\frac{\pi\ell}{m+1}\right) + \alpha \mathbf{x}_i^{\ell} \\ &= \underbrace{\left(2\beta\cos\left(\frac{\pi\ell}{m+1}\right) + \alpha\right)}_{:=\lambda_l} \mathbf{x}_i^{\ell} \quad \ell = 1,\ldots,m \end{aligned} \tag{11.2.5}$$

Somit ist gezeigt, dass $\mathbf{x}^l$ tatsächlich ein Eigenvektor von $\mathbf{M}$ ist. Als nächstes zeigen wir, dass für $\mathbf{x} = \mathbf{x}^{\ell}$ die Spalten von $\mathbf{A}$ linear abhängig sind. Einsetzen von (11.2.2) in $\mathbf{x}_{i+1} + \mathbf{x}_{i-1}$ ergibt, unter Verwendung der trigonometrischen Identitäten:

$$\begin{aligned} \mathbf{x}_{i-1}^{\ell} + \mathbf{x}_{i+1}^{\ell} = 2\cos\left(\frac{\pi l}{m+1}\right)\mathbf{x}_i^l, \quad \ell = 1,\ldots,m\ , \\ \Leftrightarrow \mathbf{A}_{:,2} = 2\cos\left(\frac{\pi l}{m+1}\right)\mathbf{A}_{:,1} \end{aligned}$$

Das bedeutet $\mathrm{Rang}(\mathbf{A}) = 1$. Daraus folgt unter Berücksichtigung von [LANM, Kor. IV.4.3.F], dass $\mathbf{A}^\top\mathbf{A}$ nicht mehr vollen Rang hat sobald $\mathbf{x}$ ein Eigenvektor von $\mathbf{M}$ ist.

Aufgabe 11.3 Parametrisierte Rang-Eins-Modifikation

Diese Aufgabe befasst sich mit speziellen Klassen von Matrizen, den symmetrischen Matrizen [LANM, Definition II.6.0.J], orthogonalen Matrizen [LANM, Definition IV.3.3.B], und sogenannten Projektionsmatrizen, für die ihr Quadrat mit ihnen selbst übereinstimmt.

In dieser Aufgabe untersuchen wir dazu Matrizen mit spezieller Struktur, welche je nach Wahl der Parameter unter Umständen zu einer der oben genannten Klassen gehören.

Wir betrachten Matrizen $\mathbf{M}$ der Form

$$\mathbf{M} = \mathbf{I}_n - \mathbf{u}\mathbf{v}^\top\ , \quad \mathbf{u},\mathbf{v} \in \mathbb{R}^n \setminus \{\mathbf{0}\}\ . \tag{11.3.1}$$

$\mathbf{M}$ ist eine sogenannnte Rang-Eins-Modifikation der Identität parametrisiert durch die Vektoren $\mathbf{u}$ und $\mathbf{v}$.

⊛ **11.3a)** Welche Bedingungen müssen $\mathbf{u}$ und $\mathbf{v}$ erfüllen, damit $\mathbf{M}$ symmetrisch ist?

Tipp: Beachten Sie, dass eine Tensorproduktmatrix [LANM, Beispiel II.3.0.H] sich nicht ändert, wenn man einen der beteiligten Vektoren mit einer Zahl $\neq 0$ multipliziert und den anderen durch die gleiche Zahl dividiert!

(→ Tipp)

Lösung: Per Definition gilt für eine symmetrische Matrix $\mathbf{M}$, $\mathbf{M} = \mathbf{M}^\top$:

$$\begin{aligned} &\mathbf{M} = \mathbf{M}^\top \\ \Leftrightarrow\ &\mathbf{u}\mathbf{v}^\top = \mathbf{v}\mathbf{u}^\top \\ \Rightarrow\ &\mathbf{u} = \xi\mathbf{v}, \quad \xi \in \mathbb{R} \end{aligned}$$

(★★) **11.3b)** Welche Bedingungen müssen $\mathbf{u}$ und $\mathbf{v}$ erfüllen, damit $\mathbf{M}$ orthogonal ist.

(→ Tipp)

Lösung:

$$\begin{aligned} &\mathbf{M}\mathbf{M}^\top = \mathbf{I} - \mathbf{u}\mathbf{v}^\top - \mathbf{v}\mathbf{u}^\top + \mathbf{u}\mathbf{v}^\top\mathbf{v}\mathbf{u}^\top = \mathbf{I} \\ \Leftrightarrow\ &-\mathbf{u}\mathbf{v}^\top - \mathbf{v}\mathbf{u}^\top + \mathbf{u}\mathbf{v}^\top\mathbf{v}\mathbf{u}^\top = 0 \end{aligned}$$

Multiplikation von (11.3b)) von rechts mit einem Vektor $\mathbf{z} \in \mathrm{Span}\{\mathbf{u}\}^\perp$ ergibt:

$$-\mathbf{u}\mathbf{v}^\top\mathbf{z} = 0,\ \forall \mathbf{z} \in \mathrm{Span}\,\{\mathbf{u}\}^\perp \tag{11.3.2}$$

Daraus schliessen wir, dass $\mathbf{v} \in \mathrm{Span}\,\{\mathbf{u}\}$ sein muss. Einsetzen von $\mathbf{v} = \xi\mathbf{u}, \xi \in \mathbb{R} \setminus 0$ ergibt:

$$\begin{aligned} -2\xi\mathbf{u}\mathbf{u}^\top + \xi^2\|\mathbf{u}\|^2\mathbf{u}\mathbf{u}^\top &= 0 \\ \xi &= \frac{2}{\|\mathbf{u}\|^2}\,. \end{aligned}$$

(★★) **11.3c)** Welche Bedingungen müssen $\mathbf{u}$ und $\mathbf{v}$ erfüllen, damit $\mathbf{M}^2 = \mathbf{M}$.

(→ Tipp)

Lösung:

$$\begin{aligned} &\mathbf{M}^2 = \mathbf{M} \\ \Leftrightarrow\ &\mathbf{I} - 2\mathbf{u}\mathbf{v}^\top + \mathbf{u}\mathbf{v}^\top\mathbf{u}\mathbf{v}^\top = \mathbf{I} - \mathbf{u}\mathbf{v}^\top \\ \Leftrightarrow\ &-\mathbf{u}\mathbf{v}^\top + \mathbf{u}\mathbf{v}^\top\mathbf{u}\mathbf{v}^\top = 0 \\ \Leftrightarrow\ &\mathbf{u}\mathbf{v}^\top(-1 + \mathbf{v}^\top\mathbf{u}) = 0 \\ \Rightarrow\ &\mathbf{v}^\top\mathbf{u} = 1 \end{aligned}$$

(★★) **11.3d)** Bestimmen Sie $\mathrm{Kern}(\mathbf{M})$ in Abhängigkeit von $\mathbf{v}$ und $\mathbf{u}$.

(→ Tipp)

Lösung: Einsetzen in die Definition von $\mathrm{Kern}(\mathbf{M})$ ergibt

$$\mathrm{Kern}(\mathbf{M}) = \{\mathbf{x} : \mathbf{M}\mathbf{x} = 0, \mathbf{x} \in \mathbb{R}^n\}$$

In Worten: Wir suchen die Menge $\{\mathbf{x}\}$, welche die Gleichung $\mathbf{M}\mathbf{x} = 0$ löst.

$$\begin{aligned} &\mathbf{M}\mathbf{x} = 0 \\ &\mathbf{x} - \mathbf{u}\mathbf{v}^\top\mathbf{x} = 0 \\ \Leftrightarrow\ &\mathbf{x} - \underbrace{(\mathbf{v}^\top\mathbf{x})}_{\in\mathbb{R}}\mathbf{u} \end{aligned}$$

Das führt uns auf den Ansatz $\mathbf{x} = \xi\mathbf{u}, \xi \in \mathbb{R}$. Einsetzen in die vorherige Gleichung liefert

$$\begin{aligned}\mathbf{u} - (\mathbf{v}^\top\mathbf{u})\mathbf{u} &= 0\\ \Rightarrow \mathbf{v}^\top\mathbf{u} &= 1\end{aligned}$$

Daraus schliessen wir, dass falls $\mathbf{v}^\top\mathbf{u} = 1$, dann gibt es ein $\mathbf{x} \in \mathbb{R}$, welches im Kern von $\mathbf{M}$ liegt.

$$\mathrm{Kern}(\mathbf{M}) = \begin{cases}\mathrm{Span}\{\mathbf{u}\} & \text{falls } \mathbf{v}^\top\mathbf{u} = 1\\ \{\mathbf{0}\} & \text{sonst}\end{cases}$$

(★★) **11.3e)** Bestimmen Sie $\mathrm{Bild}(\mathbf{M})$ für $\mathbf{u} = \mathbf{v}$ in Abhängigkeit von $\mathbf{v}$.

Tipp: Verwenden Sie das Resultat aus Teilaufgabe 11.3d).

(→ Tipp)

Lösung: Falls $\mathrm{Kern}(\mathbf{M}) = \{\}$, gilt wegen dem Dimensionssatz $\mathrm{Bild}(\mathbf{M}) = \mathbb{R}^n$. Andernfalls ist $\mathrm{Bild}(\mathbf{M})$ gegeben durch das orthogonale Komplement von $\mathrm{Kern}(M)$, also die Hyperebene mit Normalenvektor $\mathbf{u}$.

$$\begin{aligned}\mathrm{Bild}(\mathbf{M}) &= \{(\mathbf{I} - \mathbf{u}\mathbf{u}^\top)\mathbf{x} : \mathbf{x} \in \mathbb{R}^n\}\\ &= \{\mathbf{x} : \mathbf{x} \in \mathbb{R}^n, \mathbf{u}^\top\mathbf{x} = 0\}\end{aligned}$$

Aufgabe 11.4 Matrix×Vektor-Funktionen (MATLAB)

In dieser Aufgabe geht es darum zu erkennen, ob sich gewisse Abbildungen der Menge $\mathbb{R}^n$ in die Menge $\mathbb{R}^m$, welche in Form von MATLAB-Code vorliegen, durch ein Matrix-Vektor-Produkt ausdrücken lassen.

Abbildungen $\mathbb{R}^n \to \mathbb{R}^m$ werden manchmal *prozedural* beschrieben, also durch Angabe einer Funktion in einer Programmiersprache. Dann ist es oft nicht einfach zu sehen, ob sich die Abbildung als Matrix-Vektor-Produkt schreiben lässt und wie die Matrix aussieht.

In Listing 11.1 finden Sie sechs MATLAB-Funktionen, die jeweils einen Spaltenvektor `x` als Argument nehmen und einen Vektor `y` als Resultat zurückliefern. Der MATLAB-Quellcode kann auch von der Vorlesungswebseite heruntergeladen werden.

(★★) **11.4a)** Bestimmen Sie für jede Funktion den Raum (Menge) in dem der Rückgabewert `y` liegt (in Abhängigkeit von der Länge von `x`).

Tipp: Einfach mal in MATLAB ausprobieren!

(→ Tipp)

Lösung:

$f_1 \quad y \in \mathbb{R}^n$
$f_2 \quad y \in \mathbb{R}^n$
$f_3 \quad y \in \mathbb{R}^{\lfloor \frac{n}{2} \rfloor}$
$f_4 \quad y \in \mathbb{R}^n$
$f_5 \quad y \in \mathbb{R}^n$
$f_6 \quad y \in \mathbb{R}^n$

(★★) **11.4b)** Geben Sie für diejenigen Funkionen `y=f(x)`, die sich als Matrix-Vektor-Produkt $\mathbf{y} = \mathbf{A}\mathbf{x}$ schreiben lassen, die zugehörige Matrix $\mathbf{A}$ an.

(→ Tipp)

Lösung:

f_1 :

$$\mathbf{A} = \begin{bmatrix} & & & n \\ & & n-1 & \\ & \cdots & & \\ 1 & & & \end{bmatrix}$$

f_3 :

$$\mathbf{A}_{i,j} = \begin{cases} 1 & i = j \\ -1 & j = 2i \quad 1 \le j \le n, 1 \le i \le m \\ 0 & \text{sonst} \end{cases}$$

f_5 :

$$\begin{bmatrix} 2^{n-1} & 2^0 & \cdots & 2^{n-2} \\ 2^{n-2} & 2^{n-1} & \cdots & 2^{n-3} \\ \vdots & \vdots & \ddots & \vdots \\ 2^0 & 2^1 & & 2^{n-1} \end{bmatrix}$$

Listing 11.1: Verschiedene Funktionen: Matrix-Vektor-Produkt oder nicht?

```
function y = f1(x)
  y = (1:length(x))'.*x(end:-1:1);
end

function y = f2(x)
  y = x*x'*(1:length(x))';
end

function y = f3(x)
  n = length(x); y = [];
  m = floor(n/2);
  for l=1:m, y = [y;x(l) - l*x(2*l)]; end
```

```
13 end
14
15 function y = f4(x)
16   n = length(x);
17   y = (diag(x)+ones(n,1)*ones(1,n))*(n:-1:1)';
18 end
19
20 function y = f5(x)
21    n = length(x);
22    a = 2.^(0:n-1); a = [a,a,a];
23    for l=1:n
24      y(l) = 0;
25      for k=1:n, y(l) = y(l) + x(k)*a(n-l+k); end
26    end
27 end
28
29 function y = f6(x)
30   n =length(x);
31   y = (1:n)*sparse([(1:n-1)';(1:n)';(2:n)'],...
32                    [(2:n)';(1:n)';(1:n-1)'],...
33                    [x(1:n-1);x;x(2:n)],n,n);
34 end
```

Aufgabe 11.5 Effizientes Matrix-Vektor-Produkt (MATLAB)

Zur Berechnung des Matrix-Vektor-Produkts einer $n \times n$-Matrix und einem Vektor werden im Allgemeinen $O(n^2)$ Rechenoperationen benötigt [LANM, Abschnitt 5.3]. Wie Sie bereits in [LANM, Beispiel V.3.0.F] gesehen haben, lässt sich für Matrizen mit spezieller Struktur die Komplexität für das Matrix-Vektor-Produkt oft deutlich reduzieren. In dieser Aufgabe werden wir dies an einem anderen Beispiel üben.

Gegeben ist eine symmetrische Matrix $\mathbf{A} \in \mathbb{R}^{n,n}$ mit den Einträgen

$$(\mathbf{A})_{i,j} := \min(i,j) \quad i,j = 1,\ldots,n \tag{11.5.1}$$

(⋆⋆⋆) **11.5a)** Schreiben sie eine MATLAB-Funktion `Aminij_fast`, welche einen Vektor $\mathbf{x}$ entgegen nimmt und mit minimalen Rechenaufwand $\mathbf{y} = \mathbf{A}\mathbf{x}$ berechnet. Geben Sie die asymptotische Komplexität für grosse $n \in \mathbb{N}$ in der O-Notation an.

Tipp: Überlegen Sie sich zuerst die Lösung der Aufgabenstellung für den oberen Dreiecksanteil von $\mathbf{M}$. Diese Matrix entsteht aus einer oberen Dreiecksmatrix $\mathbf{D}$, deren nicht-Null-Einträge alle gleich 1 sind, durch einfache Zeilenskalierung. Die effiziente Implementierung der Multiplikation eines Vektors mit $\mathbf{D}$ ist eine einfache Summation!

(→ Tipp)

Lösung: Die Komplexität ist $O(n)$.

```
function [ y ] = Aminij(x)
%UNTITLED3 Summary of this function goes here
%   Detailed explanation goes here

n = size(x,1);

cc = (1:n)' .* flipud(cumsum(flipud(x)));
z = cumsum((1:n-1)' .* x(1:end-1));
y = cc + [0;z];

end
```

⊛ **11.5b)** Stellen Sie die Matrix $\mathbf{A}$ direkt auf. Vervollständigen Sie dazu `Aminij_direkt.m` aus den Templates.

Lösung:

```
function [ A ] = Aminij_direkt( n )
A = zeros(n);

for i = 1:n
    for j = 1:n
        A(i,j) = min(i,j);
    end
end
```

⊛ **11.5c)** Vergleichen Sie die Laufzeiten für die direkte Matrix-Vektor-Multiplikation und der schnellen Variante aus 11.5a). Verwenden Sie $n = 100, 1000, 2000, 10000$.

Tipp: Verwenden Sie das Skript `timings.m` aus den Templates.

Lösung: Die effiziente Implementierung lohnt sich, für $n = 10000$ erhält man eine Beschleunigung von ≈ 350. Siehe Abbildung 11.2.

Aufgabe 11.6 Anwendungen der vollen QR-Zerlegung (MATLAB)

In dieser Aufgabe betrachten wir einige Anwendungen der vollen QR-Zerlegung, wie sie in [LANM, Satz IV.3.4.F] eingeführt worden ist.

Die MATLAB-Funktion `qr(A)` liefert die volle QR-Zerlegung der Argumentmatrix $\mathbf{A} \in \mathbb{R}^{m,n}$, also eine *orthogonale* Matrix $\mathbf{Q} \in \mathbb{R}^{m,n}$, siehe [LANM, Definition IV.3.3.B], und eine verallgemeinerte obere Dreiecksmatrix $\mathbf{R} \in \mathbb{R}^{m,n}$ so, dass $\mathbf{A} = \mathbf{QR}$. Hat insbesondere $\mathbf{A}$ vollen Rang, so gilt ferner

$$\mathrm{Span}\{(\mathbf{A})_{:,1}, \ldots, (\mathbf{A})_{:,k}\} = \mathrm{Span}\{(\mathbf{Q})_{:,1}, \ldots, (\mathbf{Q})_{:,k}\} \quad \text{für alle} \quad k \in \{1, \ldots, \min\{m,n\}\} .$$

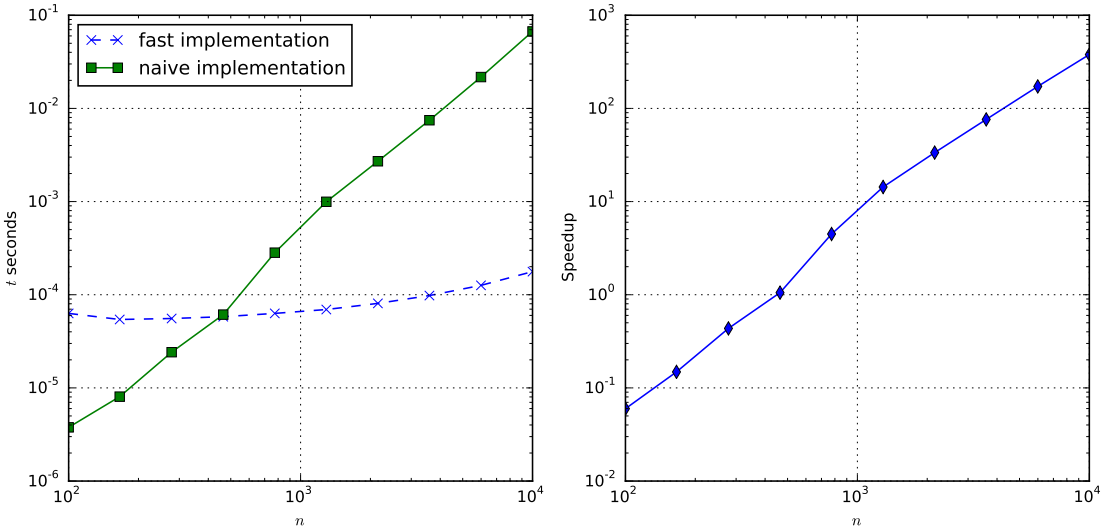

Abbildung 11.2: Plot zu 11.5c)

Dies folgt aus der Äquivalenz von QR-Zerlegung und Gram-Schmidt-Orthogonalisierung, die in [LANM, Abschnitt 4.3.4] diskutiert wurde.

In dieser Aufgabe betrachten wir die MATLAB-Funktion aus Listing 11.2 und wollen versuchen, zu verstehen, was sie leistet.

Listing 11.2: (qrgetbas.m) Eine schlecht dokumentierte MATLAB-Funktion

```
1 function B = qrgetbas(A)
2 % MATLAB function computing a matrix whose columns have a
     particular
3 % significance for the argument matrix.
4 [m,n] = size(A);
5 if (n >= m), error('A must be tall'); end
6 [Q,R] = qr(A), % full QR-decomposition
7 d = abs(diag(R)); if (min(d) < 1E-9*max(d)) % \prblabel{z7}
8     error('Inappropriate A'); end
9 B = Q(:,n+1:m);
```

⊛ **11.6a)** Wie kann man aus der vollen QR-Zerlegung einer Matrix $\mathbf{A} \in \mathbb{R}^{m,n}$, $m > n$, ablesen, dass die Matrix maximalen Rang hat?

(→ Tipp)

Lösung: $\operatorname{Rang} \mathbf{M}$ muss gleich n sein, d.h. $r_i i \neq 0, i = 1, \ldots, n$.

⊛⊛ **11.6b)** Erklären Sie die Abfrage in Zeile 7 der Funktion und schlagen Sie eine aussagekräftigere Fehlermeldung vor.

(→ Tipp)

(→ Tipp)

Lösung: Die MATLAB-Funktion `diag` gibt angwandt auf die Rechtecksmatrix $\mathbf{R}$ gibt einen Vektor mit Einträgen $r_i i, i = 1, \ldots, n$ zurück. In Zeile 7 wird die Bedingungen, die wir in 11.6a) definiert haben, überprüft. Die Fehlermeldung könnte also lauten: "Die Matrix A hat nicht vollen Rang".

(★★) **11.6c)** Was wollte wohl der Autor der MATLAB-Funktion aus Listing 11.2 berechnnen.

(→ Tipp)

(→ Tipp)

(→ Tipp)

Lösung: In Zeile 9 wird `B = Q(:, n+1:m)` gesetzt. In den Spalten von $\mathbf{Q}$ steht eine Orthonormalbasis für den $\mathbb{R}^m$, die Spalten $n+1, \ldots, m$ bilden eine Orthonormalbasis für $\mathrm{Bild}(\mathbf{A})^{\perp}$.

(★★) **11.6d)** In Teilaufgabe 11.6c) sollten Sie eine Vermutung zum Zweck der MATLAB-Funktion aus Listing 11.2 formulieren.

Implementieren Sie ein MATLAB-Skript `testqrgetbas.m`, das Ihre Vermutung für Zufallsmatrizen $\mathbf{A} \in \mathbb{R}^{9,7}$ testet.

Tipp: Vorsicht! Erinnern Sie sich an [LANM, Beispiel V.2.0.R]. Um „sauber" zu testen, ob die Differenz zweier Matrizen (bis auf Rundungsfehler) verschwindet, also ob im Rahmen der Rechengenauigkeit des Computers $\mathbf{X} - \mathbf{Y} = \mathbf{O}$, darf man keinesfalls auf Gleichheit mit $\mathbf{O}$ testen, sondern sollte etwa wie folgt vorgehen:

In MATLAB kann man die Summe der Quadrate der Einträge einer Matrix $\mathbf{M}$ so berechnen:

```
normM = sum(sum(M.*M));.
```

Um nun $\mathbf{X} - \mathbf{Y} = \mathbf{O}$ in einem numerischen Code zu testen, kann man diese Quadratsumme der Differenz nehmen und prüfen, ob sie relativ zu der entsprechenden Quadratsumme der beiden Matrizen winzig ist, also die Abfrage

```
M = X-Y;
if (sum(sum(M.*M)) < tol*max(sum(sum(X.*X)), sum(sum(Y.*Y)))
), ...,
```

wobei `tol` hinreichend klein ist, also etwa `tol = 1E-9`.

(→ Tipp)

(→ Tipp)

Lösung:

Listing 11.3: (testqrgetbas.m) Testen von `qrgetbas`

```
1 % MATLAB script for checking whether qrgetbas really computes an
2 % orthonormal basis of the orthogonal complement of the range of A
3 A = rand(9,7); % initialize random matrix
```

```
4  B = qrgetbas(A); % compute basis
5  P = B'*B;  % computer inner products of columns of B
6  M = eye(2)-P; % Check orthonormality in a safe way
7  if (sum(sum(M.*M)) < 1E-9*sum(sum(P.*P))), disp('ONB !'); end
8  M = A'*B; if (sum(sum(M.*M)) < 1E-9*sum(sum(A.*A)))
9    disp('Orthogonal complement of range A');
10 end
```

Aufgabe 11.7 QR-Zerlegung, Skalarprodukt, Orthonormalbasis

Dies Aufgabe stammt aus der Prüfung für BAUG vom Sommer 2014.

Gegeben sei die folgende MATLAB-Funktion:

```
1  function m = midvec(v1, v2, v3)
2  % Input: v1, v2, v3: linear unabhaengige Spaltenvektoren
3
4  v1 = v1/norm(v1);
5  v2 = v2/norm(v2);
6  v3 = v3/norm(v3);
7
8  w12 = v2 - v1;
9  w23 = v3 - v2;
10 w31 = v1 - v3;
11
12 [Q, R] = qr( [w12,w23,w31] );
13
14 m = Q(:, 3);
```

Im folgenden bezeichne $\mathbf{w}^{ij} \in \mathbb{R}^n$, $i,j \in \{1,2,3\}$, die Vektoren `wij` aus dem MATLAB-Code, Zeile $8-10$, und $\mathbf{v}^i \in \mathbb{R}^n$, $i = 1,2,3$, die Vektoren `vi`, die der MATLAB-Funktion als Argumente übergeben werden.

Achtung! Für die gesamte Aufgabe sei vorausgesetzt, dass $n \geq 3$ und dass die Vektoren $\mathbf{v}^1$, $\mathbf{v}^2$ und $\mathbf{v}^3$ linear unabhängig sind.

(★★) **11.7a)** Zeigen Sie, dass $\mathbf{w}^{12}$ und $\mathbf{w}^{23}$ linear unabhängig sind.

Lösung: Es ist zu zeigen, dass für $\alpha_1, \alpha_2 \in \mathbb{R}$ gilt:

$$\alpha_1 \mathbf{w}^{12} + \alpha_2 \mathbf{w}^{23} = 0 \Rightarrow \alpha_1 = \alpha_2 = 0$$

Beweis:

$$\begin{aligned} &\alpha_1 \mathbf{w}^{12} + \alpha_2 \mathbf{w}^{23} = 0 \\ \Leftrightarrow\ &\alpha_1 \left(\frac{\mathbf{v}^2}{\|\mathbf{v}^2\|} - \frac{\mathbf{v}^1}{\|\mathbf{v}^1\|} \right) + \alpha_2 \left(\frac{\mathbf{v}^3}{\|\mathbf{v}^3\|} - \frac{\mathbf{v}^2}{\|\mathbf{v}^2\|} \right) = 0 \\ \Leftrightarrow\ &-\alpha_1 \frac{\mathbf{v}^1}{\|\mathbf{v}^1\|} + (\alpha_1 - \alpha_2) \frac{\mathbf{v}^2}{\|\mathbf{v}^2\|} + \alpha_2 \frac{\mathbf{v}^3}{\|\mathbf{v}^3\|} = 0 \end{aligned}$$

$$\Rightarrow \frac{-\alpha_1}{\|\mathbf{v}^1\|} = \frac{\alpha_1 - \alpha_2}{\|\mathbf{v}^2\|} = \frac{\alpha_2}{\|\mathbf{v}^3\|} = 0 \Rightarrow \alpha_1 = \alpha_2 = 0 \qquad \square$$

(★★) **11.7b)** Es seien beliebige $i, j \in \{1, 2, 3\}$, $i \neq j$, vorgegeben. Beweisen Sie, dass gilt

$$\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\},\ \mathbf{x}^\top \mathbf{w}^{ij} = 0 \quad \Rightarrow \quad \measuredangle(\mathbf{x},\, \mathbf{v}^i) = \measuredangle(\mathbf{x},\, \mathbf{v}^j)\,.$$

In Worten, wenn ein Vektor $\mathbf{x} \in \mathbb{R}^n$ so gegeben ist, dass $\mathbf{x}^\top \mathbf{w}^{ij} = 0$ erfüllt ist, dann ist der von $\mathbf{x}$ und $\mathbf{v}^i$ gebildete Winkel gleich dem Winkel der von $\mathbf{x}$ und $\mathbf{v}^j$ gebildet wird.

Lösung: Beweis: Wir fixieren beliebige $i, j \in \{1, 2, 3\}, i \neq j$ und nehmen im $\mathbf{x} \in \mathbb{R}^n \setminus \{0\}$, so dass $\mathbf{x}^\top \mathbf{w}^{ij} = 0$. Es gilt

$$\begin{aligned} 0 = \mathbf{x}^\top \mathbf{w}^{ij} = \langle \mathbf{w}^{ij}, \mathbf{x} \rangle &= \langle \frac{\mathbf{v}^j}{\|\mathbf{v}^j\|} - \frac{\mathbf{v}^i}{\|\mathbf{v}^i\|}, \mathbf{x} \rangle \\ \Leftrightarrow \frac{\|\mathbf{v}^j\|}{\|\mathbf{v}^j\|}\|\mathbf{x}\|\measuredangle(\mathbf{v}^j, \mathbf{x}) &= \frac{\|\mathbf{v}^i\|}{\|\mathbf{v}^i\|}\|\mathbf{x}\|\measuredangle(\mathbf{v}^i, \mathbf{x}) \\ \Leftrightarrow \measuredangle(\mathbf{v}^j, \mathbf{x}) &= \measuredangle(\mathbf{v}^i, \mathbf{x}) \qquad \square \end{aligned}$$

(★★) **11.7c)** Es stehe $\mathbf{Q} \in \mathbb{R}^{n,n}$ für die in Zeile 12 des obigen MATLAB-Codes durch Aufruf der MATLAB-Funktion `qr` berechnete Matrix. Zeigen Sie, dass die ersten beiden Spalten von $\mathbf{Q}$ den gleichen Raum aufspannen wie $\mathbf{w}^{12}$ und $\mathbf{w}^{23}$. ($\rightarrow$ Tipp)

Lösung: Es gilt $\mathbf{A} = \mathbf{QR}$, in der ersten Spalte von $\mathbf{A}$ steht $\mathbf{w}^{12}$, in der zweiten Spalte von $\mathbf{A}$ steht $\mathbf{w}^{23}$. Wir benutzen die obere Dreiecksform von $\mathbf{R}$.

$$\begin{aligned} \mathbf{w}^{12} &= \mathbf{R}_{1,1}\mathbf{Q}_{:,1} = (\mathbf{QR})_{:,1} \\ \mathbf{w}^{23} &= \mathbf{R}_{1,2}\mathbf{Q}_{:,1} + \mathbf{R}_{2,2}\mathbf{Q}_{:,2} = (\mathbf{QR})_{:,2} \end{aligned}$$

Damit ist die Behauptung gezeigt.

(★★) **11.7d)** Sei $\mathbf{m} \in \mathbb{R}^n$ der Ausgabevektor `m` der oben angegebenen MATLAB-Funktion. Zeigen Sie, dass $\mathbf{m}$ mit allen Vektoren $\mathbf{v}^1,\ \mathbf{v}^2,\ \mathbf{v}^3$, den gleichen Winkel einschliesst, das heisst

$$\measuredangle(\mathbf{m},\, \mathbf{v}^1) = \measuredangle(\mathbf{m},\, \mathbf{v}^2) = \measuredangle(\mathbf{m},\, \mathbf{v}^3)\,.$$

Tipp: Die Aussage kann durch Kombination der Aussagen vorhergehender Teilaufgaben bewiesen werden.

Lösung: $\mathbf{m} := \mathbf{q}^3$, wenn wir die Notation aus 11.7c) weiter verwenden. Per Definition von $\mathbf{Q}$ gilt:

$$\begin{aligned} \mathbf{m}^\top \mathbf{q}^1 = \langle \mathbf{q}^1, \mathbf{m} \rangle = 0,\ \mathbf{m}^\top \mathbf{q}^2 = \langle \mathbf{q}^2, \mathbf{m} \rangle = 0 \\ \Rightarrow \mathbf{m}^\top(\alpha_1 \mathbf{q}^1 + \alpha_2 \mathbf{q}^2) = 0 \quad \forall \alpha_1, \alpha_2 \in \mathbb{R} \\ \Rightarrow \mathbf{m}^\top \mathbf{u} \quad \forall \mathbf{u} \in \operatorname{Span}\{\mathbf{q}^1, \mathbf{q}^2\} = \operatorname{Span}\{\mathbf{w}^{12}, \mathbf{w}^{23}\} \\ \Rightarrow \measuredangle(\mathbf{m}, \mathbf{v}^1) = \measuredangle(\mathbf{m}, \mathbf{v}^2) = \measuredangle(\mathbf{m}, \mathbf{v}^3) \qquad \square \end{aligned}$$

Tipps

Tipps für Aufgabe 11.1

11.1b) Betrachten Sie die zweite Gleichung von (11.1.1).

11.1c) In [LANM, Abschnitt 2.7] haben Sie die Blockelimination für lineare Gleichungssysteme kennengelernt. Wenden Sie diese nun auf Gleichung (11.1.1) an und erhalten Sie dadurch die Normalengleichungen als *Schur-Komplement*, siehe [LANM, Gleichung II.7.0.F].

11.1d) Denken Sie an das Problem aus Aufgabe 10.9 (Teilaufgaben 10.9d) und 10.9e)) von Serie 10 für eine sehr grosse Anzahl von Messpunkten. Dann erhalten wir eine *grosse* und *dünnbesetzte* Matrix $\mathbf{A}$. Was kann passieren, wenn man in diesem Fall die Normalengleichungsmatrix $\mathbf{A}^\top\mathbf{A}$ aufstellt?

Tipps für Aufgabe 11.2

11.2a) Die Herausforderung besteht darin, die Koeffizientenmatrix $\mathbf{A}$, die Unbekannten $\mathbf{x}$ und die rechte Seite $\mathbf{b}$ aus (11.2.1) richtig zu identifizieren. Dabei ist Vorsicht geboten, denn die Notation in (11.2.1) kann leicht in die Irre führen; die $\mathbf{x}_i$ sind keine Unbekannten!

11.2b) Verwenden Sie das Resultat von 11.2a). Mit Hilfe des „Matrixbaukastens" von MATLAB können Sie dann die Koeffizientenmatrix des überbestimmten linearen GLeichungssystems initialisieren und die Kleinste-Quadrate-Lösung bestimmen, wie in [LANM, Abschnitt 5.5] beschrieben.

11.2d) Erinnern Sie sich an [LANM, Kor. IV.4.3.F], das eine Bedingung an die Systemmatrix für die Eindeutigkeit der Kleinste-Quadrate-Lösung gibt.

Überlegen Sie sich dann, wann die Matrix des überbestimmten Gleichungssystems aus Teilaufgabe 11.2a) keine vollen Spaltenrang hat. Dies führt auf ein Eigenwertproblem, wie in [LANM, Abschnitt 7.4] der Vorlesung besprochen.

Um Eigenvektoren zu finden probieren Sie schliesslich Eingangssignale der Form

$$\mathbf{x}^\ell = (x_j^\ell)_{j=1}^m \in \mathbb{R}^m\,, \quad \ell \in \{1,\ldots,m\} \quad \text{mit} \quad x_j^\ell := \sin(\pi\frac{j\ell}{m+1})\,, \quad j = 1,\ldots,m\,.$$

Tipps für Aufgabe 11.3

11.3a) Leiten Sie aus $\mathbf{M} = \mathbf{M}^\top$ eine Gleichheit zweier Tensorproduktmatrizen ab. Verwenden Sie dann, dass zwei Matrizen nur dann gleich sein können, wenn auch ihre Bilder übereinstimmen. Was ist $\mathrm{Bild}(\mathbf{T})$ für eine Tensorproduktmatrix $\mathbf{T}$?

11.3b) Starten Sie von $\mathbf{M}\mathbf{M}^\top = \mathbf{I}$, setzen Sie die angegebene Formel für MM ein, multiplizieren Sie aus, und kürzen Sie $\mathbf{I}$ auf beiden Seiten. Multiplizieren Sie die erhaltenen Matrixindentität auf beiden Seiten von rechts mit einem zu $\mathbf{u}$ orthogonalen Spaltenvektor $\mathbf{z} \in \mathrm{Span}\{\mathbf{u}\}^\perp$. Was muss dieser Vektor in Bezug auf $\mathbf{v}$ erfüllen? Denken Sie dann

daran, dass Ihre Beobachtung für *jeden* Vektor aus $\mathrm{Span}\{\mathbf{u}\}^\perp$ gelten muss. Schliesslich können Sie noch benutzen, dass $\left(\mathrm{Span}\{\mathbf{u}\}^\perp\right)^\perp = \mathrm{Span}\{\mathbf{u}\}$.

11.3c) Wiederum geht man von der Matrixidentität $\mathbf{M}^2 = \mathbf{M}$ aus, setzt ein, multipliziert aus und kürzt. Nutzen Sie dann aus, dass ein inneres Produkt zweier Vektoren eine Zahl ergibt, die man aus dem Matrixprodukt herausziehen kann.

11.3d) Welche Gleichung muss $\mathbf{x} \in \mathrm{Kern}(\mathbf{M})$ für ein $\mathbf{M}$ der angegebenen Form erfüllen.

Mit dieser Vorüberlegung können Sie dann folgende Frage beantworten: Von welchem Vektor muss $\mathrm{Kern}(\mathbf{M})$ notwendig erzeugt werden, wenn $\mathrm{Kern}(\mathbf{M}) \neq \{\mathbf{0}\}$?

Wann kann also $\mathrm{Kern}(\mathbf{M}) \neq \{\mathbf{0}\}$ überhaupt nur eintreten? Führen Sie eine Fallunterscheidung durch.

11.3e) Zusammen mit dem Resultat aus Teilaufgabe 11.3d) kann hier der Dimenssionssatz [LANM, Thm. III.3.0.H] angewandt werden. Wenn $\mathrm{Kern}(\mathbf{M}) = \{\mathbf{0}\}$ dann liefert er die Lösung.

Sonst kann man die Symmetrie von $\mathbf{M}$ für $\mathbf{v} = \mathbf{u}$ ausnutzen, die impliziert, dass dann $\mathrm{Bild}(\mathbf{M}) = \mathrm{Kern}(\mathbf{M})^\perp$.

Tipps für Aufgabe 11.4

11.4a)
- MATLAB legt mit `(i:j)` einen *Zeilenvektor* an.
- Mit dem `;`-Operator kann man in MATLAB Komponente um Komponente an einen Spaltenvektor anfügen.

11.4b) Die Matrix kann man sich für eine vorgegebene Länge von `x` leicht von MATLAB ausrechnen lassen, wenn man berücksichtigt, dass in der j. Spalte der Abbildungsmatrix einfach das Bild des j. Einheitsvektors stehen muss, wie in [LANM, Abschnitt 6.3] der Vorlesung erklärt.

Tipps für Aufgabe 11.5

11.5a) Die MATLAB-Funktion `cumsum` könnte hilfreich sein. Informieren Sie sich mittels dem Kommando `help cumsum`.

Tipps für Aufgabe 11.6

11.6a) Erinnern Sie sich an [LANM, Satz III.3.0.I], nämlich die Invarianz des Rangs einer Matrix bei Multiplikation mit einer invertierbaren Matrix. Orthogonale Matrizen sind per definitionem invertierbar.

11.6b) Hier sollen Sie natürlich das Erkenntnis aus Teilaufgabe 11.7a) benutzen. Stützen Sie sich auf die Hilfe-Funktion von MATLAB, um in Erfahrung zu bringen, was diag angewandt auf eine Rechtecksmatrix zurückgibt.

11.6b) Um die Abfrage in Zeile 7 von 11.2 zu verstehen, muss man sich auch an die Diskussion der Abbruchbedingung aus [LANM, Listing 5.4] in [LANM, Abschnitt 5.2] erinnern.

11.6c) Überlegen Sie sich, was die Menge $\{(\mathbf{B})_{:,1}, \ldots, (\mathbf{B})_{:,m-n}\}$ repräsentiert.

11.6c) Wenn $\mathbf{Q} \in \mathbb{R}^{n,n}$ eine orthogonale Matrix ist, was bilden die Spalten $(\mathbf{Q})_{:,k+1}, \ldots, (\mathbf{Q})_{:,n}$ in Bezug auf den Span der Spalten $(\mathbf{Q})_{:,1}, \ldots, (\mathbf{Q})_{:,k}$? Erinnern Sie sich an den Begriff des *orthogonalen Komplements* aus [LANM, Definition IV.3.2.B].

11.6c) Um den Zweck von `qrgetbas` zu beschreiben ist der Begriff der *Orthonormalbasis* (ONB) aus [LANM, Definition IV.3.1.A] nützlich.

11.6d) Erinnern Sie sich daran, wie wir in [LANM, Listing 5.6] die Orthonormalität von Vektoren getestet haben, die als Spalten einer Matrix gegeben waren.

11.6d) Wenn $\mathbf{AB} = \mathbf{O}$, dann wissen wir, dass die Spalten von $\mathbf{B}$ senkrecht auf $\mathrm{Bild}(\mathbf{A})$ stehen.

Tipps für Aufgabe 11.7

11.7c) Erinnern Sie sich daran, was

```
[Q, R] = qr(A);
```

bedeutet: Für $\mathbf{A} \in \mathbb{R}^{n,m}$, mit $n \geq m$, werden zwei Matrizen $\mathbf{Q} \in \mathbb{R}^{n,n}$ und $\mathbf{R} \in \mathbb{R}^{n,m}$ berechnet, so dass

- $\mathbf{QR} = \mathbf{A}$,
- $\mathbf{Q} \in \mathbb{R}^{n,n}$ ist eine *orthogonale Matrix*,
- $\mathbf{R}$ ist eine *obere Dreiecksmatrix*.
- Drücken Sie die Matrix $\mathbf{Q}$ anhand ihrer Spaltenvektoren

$$\mathbf{q}^1, \ldots, \mathbf{q}^n \in \mathbb{R}^n$$

aus und berechnen Sie damit die ersten beiden Spalten des Produktes $\mathbf{QR}$, um damit die Behauptung zu beweisen.

Veröffentlichung am 10. Dezember 2014.
Abzugeben bis 17. Dezember 2014.

Literatur

[LANM] Vorlesungszusammenfassung für die Vorlesung "Lineare Algebra und Numerische Mathematik, D-BAUG".

Last modified on 13. Januar 2015