

Deterministische Signalverarbeitung

Lineare Operatoren

Univ.-Prof.,Dr.-Ing. Markus Rupp
SS 04

Letzte Änderung: 13.05.2004

Lernziele

Lineare Operatoren

- Lineare Transformationen, Funktionale
- Null- und andere Räume
- Orthog. Unterräume, Matrix Rang
- Projektionen, Projektionssatz
- Faktorisierung
 - Eigenwertzerlegung
 - Filterberechnung basierend auf Eigenfilter
- Unterraumtechniken:
 - PHD,MUSIC,ESPRIT
 - SVD
 - Konditionszahl

Resumee: Vektorräume

- **Definition 2.11:** Ein linearer Vektorraum S über einer Menge von Skalaren $T (C,R,Q,Z,N,B)$ ist eine Ansammlung von Vektoren zusammen mit einer additiven und multiplikativen Operation, die folgende Eigenschaften erfüllen:
- 1) S formt eine Gruppe bzgl. der Addition
- 2) Für jedes $a,b \in T$ und x,y in S gilt:
 $a\underline{x} \in S, a(b\underline{x}) = (ab)\underline{x}, (a+b)\underline{x} = a\underline{x} + b\underline{x}$
 $a(\underline{x} + \underline{y}) = a\underline{x} + a\underline{y}$
- 3) Bzgl. der Multiplikation existiert ein Eins- und ein Nullelement:

$$1\underline{x} = \underline{x}; \quad 0\underline{x} = \underline{0}$$

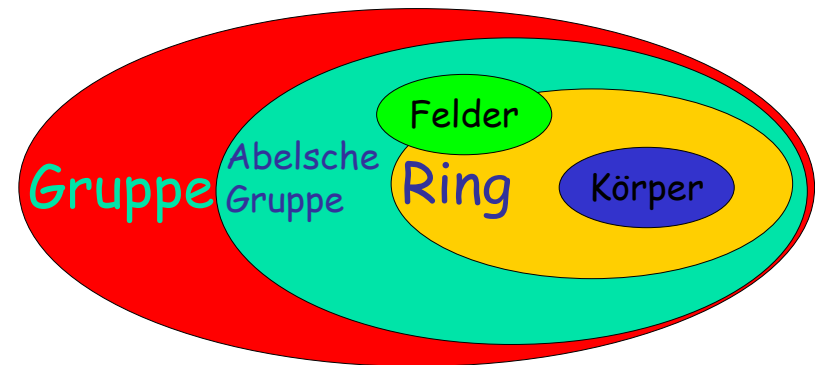
Resumee: Gruppen

- **Definition 2.8 (Gruppe):** Eine Menge S , für die eine binäre Operation $\star$ (Operation bzgl. zweier Elemente von S) definiert ist, heißt Gruppe, wenn gilt:
 - 1) Für jedes a,b in S gilt: $(a \star b)$ in S
 - 2) es existiert ein Identitätselement e in S , so dass für jedes Element a in S gilt: $e \star a = a \star e = a$.
 - 3) Für jedes Element a in S existiert ein inverses Element b in S , so dass: $a \star b = b \star a = e$.
 - 4) Die binäre Operation $\star$ ist assoziativ, d.h.: $(a \star b) \star c = a \star (b \star c)$
- Die Gruppe wird mit $(S, \star)$ bezeichnet.
 - Gilt weiterhin für jedes Paar a,b aus S , dass $a \star b = b \star a$ (kommutativität), so heißt die Gruppe kommutativ oder Abelsch.

Resumee: Ringe

- **Definition 2.9 (Ring):** Eine Menge S , für die zwei binäre Operationen $+$ und $\star$ definiert sind, heißt Ring, wenn folgendes gilt:
 - $(S,+)$ ist eine Abelsche Gruppe
 - Die Operation $\star$ ist assoziativ.
 - Es gilt Distributivität bzgl. $+$: $a(b+c)=ab+ac$, $(a+b)c=ac+bc$.
- Ein Ring wird mit $(S,+, \star)$ bezeichnet.
- Beachte: für $\star$ muss es kein Identitäts oder inverses Element geben.
- Existiert zusätzlich das inverse Element zu $\star$, dann spricht man von einem Körper.

Resumee: Körper



Linearität

- **Definition 4.1:** Eine Transformation $A: X \rightarrow Y$ in denen X und Y Vektorräume sind, die über einem Ring R definiert sind, heißt linear, wenn für jedes x_1, x_2 aus X und Faktoren α_1, α_2 aus R gilt:

$$A(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 A(x_1) + \alpha_2 A(x_2)$$
- Beispiele für lineare Operatoren sind Matrizen, Abtastung, Ableitungen und Faltungsintegrale (Funktionale)

Linearität

- Beispiele für lineare Operatoren sind:
 - **Beispiel 4.1** Eine komplex-wertige Zahl z aus C werde in einen Vektor $\underline{x}$ aus R^2 abgebildet:

$$Az = \underline{x} = [\text{real}(z), \text{imag}(z)]^T$$
 - **Beispiel 4.2** Eine Quadrupel $\{s_1, s_2, s_3, s_4\}$ aus C werde in eine 4×4 Matrix aus C abgebildet:

$$A\{s_1, s_2, s_3, s_4\} = \begin{bmatrix} s_1 & s_2 & s_3 & s_4 \\ -s_2^* & s_1^* & -s_4^* & s_3^* \\ s_3^* & s_4^* & s_1^* & s_2^* \\ s_4 & -s_3 & s_2 & -s_1 \end{bmatrix}$$

Linearität

- Beispiele für lineare Operatoren sind:
 - Beispiel 4.3** Eine kontinuierliche Funktion $g(t)$ aus $C[0,1]$ werde zu festen Zeitpunkten $0 < t_1 < t_2 < \dots < t_n < 1$ abgetastet:
 $Ag(t) = [g(t_1), g(t_2), \dots, g(t_n)]$
 - Beispiel 4.4:** Funktionen, die eine Menge X auf reelle Zahlen abbildet, heißen Funktionale. Sind diese linear, so heißen sie lineare Funktionale:

$$f_1(x) = \frac{1}{T} \int_0^T x(t) dt$$

$$f_2(x) = \int_a^b x(t) g(\tau - t) dt$$

$$f_3(x) = \int_{-\infty}^{\infty} x(t) \exp(-j\omega t) dt$$

Beschränkte Operatoren

- Wir hatten bereits Vektornormen eingeführt und gezeigt, dass diese als induzierte Normen ebenso für Matrizen verwendet werden können:

$$\|A\|_p = \|A\|_{p, \text{ind}} = \sup_{x \neq 0} \frac{\|Ax\|_p}{\|x\|_p} = \sup_{x \neq 0} \left\| A \left(\frac{x}{\|x\|_p} \right) \right\|_p = \sup_{\|x\|_p=1} \|Ax\|_p$$

- Definition 4.2:** Ist die Norm eines Operators endlich, so heißt der Operator beschränkt (engl.: bounded).

Beschränkte Operatoren

- Satz 4.1:** Ein linearer Operator $A: X \rightarrow Y$ ist dann und nur dann beschränkt, wenn er kontinuierlich ist.
- Beweis:** Nehmen wir an, dass A beschränkt ist, so dass $\|Ax\| \leq M \|x\|$ für alle x aus X . Dann muss dies auch für alle konvergenten Folgen x_n aus X gelten:

$$\|Ax_n\| \leq M \|x_n\| \rightarrow 0$$

- Da kontinuierliche Funktionen Konvergenz bewahren, folgt dass auch A kontinuierlich sein muss.

Beschränkte Operatoren

- Nun zum Gegenteil: Nehmen wir an, dass A kontinuierlich ist. Dann muss es für ein $\delta > 0$ ein $\|x\| < \delta$ geben, so dass $\|Ax\| < 1$. Somit gilt:

$$\|Ax\| = \underbrace{\|A(\delta x / \|x\|)\|}_{< 1} \|x\| / \delta < \|x\| / \delta$$

- Damit ist $M = 1/\delta$ die obere Schranke für A .
- Dieser Satz gilt für lineare Operatoren, also auch für lineare Funktionale.

Innere Produkte

- **Lemma 4.1:** Innere Produkte sind kontinuierlich. D.h. wenn $\underline{x}_n \rightarrow \underline{x}$ in einem inneren Produktraum S gilt, dann gilt auch $\langle \underline{x}_n, \underline{y} \rangle \rightarrow \langle \underline{x}, \underline{y} \rangle$ für $\underline{y}$ aus S .

- **Beweis:** Wenn $\underline{x}_n$ konvergiert, muss es auch beschränkt sein, also $\|\underline{x}_n\| \leq M < \infty$

- Dann gilt auch:

$$\|\langle \underline{x}_n, \underline{y} \rangle - \langle \underline{x}, \underline{y} \rangle\| = \|\langle \underline{x}_n - \underline{x}, \underline{y} \rangle\| \leq \|\underline{x}_n - \underline{x}\| \|\underline{y}\|$$

Da $\|\underline{x}_n - \underline{x}\| \rightarrow 0$, konvergiert $\langle \underline{x}_n, \underline{y} \rangle$ gegen $\langle \underline{x}, \underline{y} \rangle$.

Beschränkte Operatoren

- Nicht unwichtig sind solche Eigenschaften auch für die Inversen linearer Operatoren.

- **Satz 4.2:** Sei $\|\cdot\|$ eine Matrixnorm, welche die submultiplikative Eigenschaft erfüllt und $A: X \rightarrow X$ ein linearer Operator mit $\|A\| < 1$. Dann existiert $(I-A)^{-1}$ und:

$$(I-A)^{-1} = \sum_{i=0}^{\infty} A^i$$

Beschränkte Operatoren

- **Beweis:** Sei $\|A\| < 1$. Wenn $I-A$ singular ist, dann gibt es mindestens einen Vektor $\underline{x}$ ungleich $\underline{0}$, so dass $(I-A)\underline{x} = \underline{0}$. Damit gilt aber $\underline{x} = A\underline{x}$ und somit

$$\|\underline{x}\| = \|A\underline{x}\| \leq \|\underline{x}\| \|A\|$$

- Dann muss aber $\|A\| > 1$ sein, was einen Widerspruch darstellt.
- Durch Multiplikation erhält man:

$$(I-A)(I + A + A^2 + \dots + A^{k-1}) = I - A^k$$

Beschränkte Operatoren

- Da $\|A^k\| \leq \|A\|^k$
- und $\|A\| < 1$ muss gelten, dass

$$\lim_{k \rightarrow \infty} \|A\|^k = 0$$

- Somit gilt auch:

$$(I-A) \left(\sum_{i=0}^{\infty} A^i \right) = I$$

Anmerkung

- Da die meisten linearen Operationen durch Matrizen beschrieben werden können, werden wir uns im folgenden auch hauptsächlich mit linearen Abbildungen in Form von Matrizen befassen.
- Viele der gezeigten Eigenschaften insbesondere die Projektionseigenschaften sind aber nicht auf Matrizen beschränkt!

Null- und andere Räume

- **Definition 4.3:** Der Vektorraum, der durch die Spalten einer Matrix $A=[\underline{a}_1, \underline{a}_2, \dots, \underline{a}_n]: X \rightarrow Y$ aufgespannt wird, heißt Spaltenraum von A (engl.: column space or range):

$$R(A) = \text{span}(\underline{a}_1, \underline{a}_2, \dots, \underline{a}_n) \\ = \{ \underline{y} \in Y : A\underline{x} = \underline{y} \text{ für } \underline{x} \in X \}$$

- Die zweite Zeile ist die allgemeine Form und beschreibt den Spaltenraum eines linearen Operators.

Null- und andere Räume

- Sei ein Vektor $\underline{b}$ aus dem Spaltenraum von A , dann gilt:
Eine Linearkombination der Spalten von A muss exakt $\underline{b}$ sein: $A\underline{x} = \underline{b}$.
- Gegeben $A\underline{x} = \underline{b}$,
 - Es existiert genau eine Lösung, wenn $\underline{b}$ im Spaltenraum von A liegt und die Spalten linear unabhängig sind.
 - Es existiert keine Lösung, wenn $\underline{b}$ nicht im Spaltenraum von A liegt.
 - Es existieren unendlich viele Lösungen, wenn $\underline{b}$ im Spaltenraum von A liegt, die Spalten aber nicht linear unabhängig sind.
 - Beweise folgen später...

Null- und andere Räume

- **Definition 4.4:** Der Vektorraum, der durch die (konjugiert komplexen) Zeilen einer Matrix $A=[\underline{b}_1^T; \underline{b}_2^T; \dots; \underline{b}_n^T]: X \rightarrow Y$ aufgespannt wird, heißt Zeilenraum von A (engl.: row space) oder Spaltenraum der Adjunkten Matrix A^H :

$$R(A^H) = \text{span}(\underline{b}_1^H; \underline{b}_2^H; \dots; \underline{b}_n^H) \\ = \{ \underline{x} \in X : A^H \underline{y} = \underline{x} \text{ für } \underline{y} \in Y \}$$

Null- und andere Räume

- **Definition 4.5:** Der Vektorraum, der durch die Lösungen $Ax=0$ eines linearen Operators $A:X \rightarrow Y$ gekennzeichnet ist, heißt Nullraum $N(A)$ von A (engl.: nullspace). Er wird auch Kern (engl.: kernel) des Operators genannt.
- **Definition 4.6:** Der Vektorraum, der durch die Lösungen $A^H y=0$ eines linearen Operators $A:X \rightarrow Y$ gekennzeichnet ist, heißt Nullraum $N(A^H)$ von A^H oder auch linker Nullraum von A (engl.: left nullspace).

Null- und andere Räume

- **Beispiel 4.5:** Sei A ein linearer Operator mit:

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

- Dann sind der Spalten- und der Nullraum von A gegeben durch:

$$R(A) = \text{span} \left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right); \quad N(A) = \text{span} \left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right)$$

Null- und andere Räume

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

- **Noch Beispiel 4.5:** Der Zeilenraum, bzw. der Spaltenraum der Adjunkten Matrix und der linke Nullraum sind gegeben durch:

$$R(A^H) = \text{span} \left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right); \quad N(A^H) = \text{span} \left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right)$$

Null- und andere Räume

- **Beispiel 4.6:** Betrachte die Faltungsoperation:

$$L\{x(t)\} = \int_0^t x(\tau)h(t-\tau)d\tau$$

- Nullraum des linearen Operators L besteht aus den Funktionen $x(t)$, die gefaltet mit $h(t)$ Null ergeben.
- Im Fourierbereich entspricht dies den Funktionen $X(j\omega)$, die keinen gemeinsamen Wertebereich mit $H(j\omega)$ haben. Also:

$$\{x(t); H(j\omega)X(j\omega) \equiv 0\}$$

Null- und andere Räume

- Aufgrund ihrer Definitionen können wir bereits folgende Beziehungen für lineare Operatoren $A: X \rightarrow Y$ angeben:

$$R(A) \subset Y$$

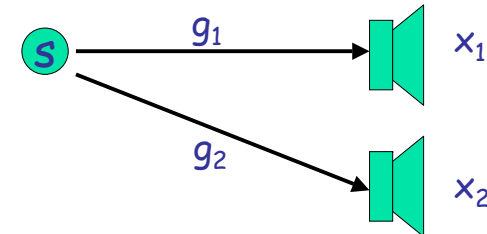
$$R(A^H) \subset X$$

$$N(A) \subset X$$

$$N(A^H) \subset Y$$

Beispiel 4.7

- Betrachte das Problem der Stereoübertragung einer Quelle s :



- Hinweis: g_i seien die unterschiedlichen Wege zu den Mikrofonen.

Beispiel 4.7

- Mit der Vektornotation finden wir folgende Zusammenhänge:

$$\underline{x}_{i,k} = [x_i(k), x_i(k-1), \dots, x_i(k-L+1)]^T$$

$$\underline{g}_i = [g_{i,0}, g_{i,1}, \dots, g_{i,L-1}]^T$$

$$S_k = \begin{bmatrix} s(k) & s(k-1) & \dots & s(k-L+1) \\ s(k-1) & s(k-1) & & \\ \vdots & & \ddots & \\ s(k-L+1) & & & s(k-2L+2) \end{bmatrix} = S^T$$

$$\underline{x}_{i,k} = S_k \underline{g}_i; \quad i = 1, 2$$

Beispiel 4.7

- Beachte folgenden Zusammenhang:

$$\begin{aligned} \underline{x}_{i,k}^T \underline{g}_j &= \underline{g}_i^T S^T \underline{g}_j = \underline{g}_i^T S \underline{g}_j \\ &= \underline{g}_j^T S \underline{g}_i = \underline{x}_{j,k}^T \underline{g}_i \end{aligned}$$

- Bilde den Vektor $\underline{x}_k^T = [x_{1,k}^T, x_{2,k}^T]$.
- Damit ist aber die folgende AKF Matrix singular!

$$R_{\underline{xx}}(n) = \sum_{k=0}^n \underline{x}_k \underline{x}_k^T = \sum_{k=0}^n \begin{pmatrix} S_k \underline{g}_1 \underline{g}_1^T S_k & S_k \underline{g}_1 \underline{g}_2^T S_k \\ S_k \underline{g}_2 \underline{g}_1^T S_k & S_k \underline{g}_2 \underline{g}_2^T S_k \end{pmatrix}$$

Beispiel 4.7

- Bilde folgenden Vektor für beliebige Werte α_i :

$$\underline{u}^T = [\alpha_2 \underline{g}_2^T, -\alpha_2 \underline{g}_1^T]$$

- Für diesen Vektor gilt:

$$R_{xx}(n)\underline{u} = \underline{0}$$

- Offensichtlich liegt $\underline{u}$ im Nullraum von $R_{xx}(n)$.

Orthogonale Unterräume

- Definition 4.7:** Sei S ein Vektorraum und V und W beide Unterräume von S . V und W heißen orthogonale Unterräume wenn für jedes Paar von Vektoren $\underline{v}$ aus V und $\underline{w}$ aus W gilt: $\langle \underline{v}, \underline{w} \rangle = 0$.
- Definition 4.8:** Sei V eine Teilmenge eines Vektorraumes S mit innerem Vektorprodukt. Der Raum aller Vektoren, die orthogonal zu den Vektoren in V sind, heißt orthogonaler Komplementärraum von V (engl.: orthogonal complement) und wird folgendermaßen bezeichnet:

$$V^\perp$$

Orthogonale Unterräume

- Beispiel 4.8:** Sei $S = GF(2)^3$. Die Vektoren $\underline{v} = (1, 0, 0)$ und $\underline{w} = (0, 0, 1)$ seien aus S . Sie spannen die Unterräume V und W auf:
 $V = \text{span}(\underline{v}) = \{(0, 0, 0), (1, 0, 0)\}$
 $W = \text{span}(\underline{w}) = \{(0, 0, 0), (0, 0, 1)\}$
- Beide Räume sind orthogonale Unterräume zueinander.
- Der Unterraum V hat den orthogonalen Komplementärraum:
 $V^\perp = \{(0, 0, 0), (0, 1, 0), (0, 0, 1), (0, 1, 1)\}$
- Von welchen Vektoren wird der orthogonale Komplementärraum aufgespannt?

Orthogonale Unterräume

- Satz 4.3:** Seien V und W zwei Unterräume eines Vektorraums S (der nicht notwendigerweise abgeschlossen ist) mit innerem Vektorprodukt. Dann gilt:
 - $V^\perp$ ist ein abgeschlossener Unterraum von S
 - $V \subset V^{\perp\perp}$
 - Wenn $V \subset W$, dann gilt: $W^\perp \subset V^\perp$
 - $V^{\perp\perp} = V^\perp$
 - Wenn $\underline{x} \in V \cap V^\perp$, dann $\underline{x} = \underline{0}$
 - $\{\underline{0}\}^\perp = S, S^\perp = \{\underline{0}\}$

Orthogonale Unterräume

- **Beweis (Teil 1):** Sei $\underline{x}_n$ eine Folge von Vektoren in $V^\perp$ mit $\underline{x}_n \rightarrow \underline{x}$. Damit der Raum abgeschlossen ist, muss der Grenzwert $\underline{x}$ auch in $V^\perp$ liegen. Wegen dem inneren Vektorprodukt (Lemma 4.1) gilt:

$$\langle \underline{x}_n, \underline{v} \rangle = 0 \Rightarrow \langle \underline{x}, \underline{v} \rangle = 0; \text{ für } \underline{v} \in V \\ \Rightarrow \underline{x} \in V^\perp$$

- **Hinweis:** Beachte, dass nicht gilt: $V = V^{\perp\perp}$. Dies liegt an der fehlenden Abgeschlossenheit des Raumes S . Es mag Cauchy-Folgen in V geben, die nicht in $V^{\perp\perp}$ liegen.

Orthogonale Unterräume

- **Satz 4.4:** Sei $A: X \rightarrow Y$ ein beschränkter, linearer Operator zwischen zwei Hilberträumen X und Y und $R(A)$ sowie $R(A^H)$ abgeschlossen, so gilt:

- 1) $\{R(A)\}^\perp = N(A^H); \quad R(A) = \{N(A^H)\}^\perp$
- 2) $\{R(A^H)\}^\perp = N(A); \quad R(A^H) = \{N(A)\}^\perp$

- **Beweis:** Sei $\underline{z} \in N(A^H)$ und $\underline{y} \in R(A)$
mit $A\underline{x} = \underline{y}$ für $\underline{x} \in X$
 $\langle \underline{y}, \underline{z} \rangle = \langle A\underline{x}, \underline{z} \rangle = \langle \underline{x}, A^H \underline{z} \rangle = \langle \underline{x}, 0 \rangle = 0$
Also muss $\underline{y} \perp \underline{z}$ und somit $N(A^H) \subset \{R(A)\}^\perp$

Orthogonale Unterräume

- Weiter: $\underline{y} \in \{R(A^H)\}^\perp$ mit $A\underline{x} = \underline{y}$ für $\underline{x} \in X$
 $\langle A\underline{x}, \underline{y} \rangle = 0 = \langle \underline{x}, A^H \underline{y} \rangle$
Also muss $A^H \underline{y} = 0$ und somit $\underline{y} \in N(A^H)$,
also $\{R(A)\}^\perp \subset N(A^H)$
- Da also $N(A^H) \subset \{R(A)\}^\perp$ und $\{R(A)\}^\perp \subset N(A^H)$
muss gelten: $N(A^H) = \{R(A)\}^\perp$

Orthogonale Unterräume

- **Satz 4.5:** (Fredholm'scher Alternativsatz) Sei A ein beschränkter Operator. Die Gleichung $A\underline{x} = \underline{b}$ hat dann und nur dann eine Lösung wenn $\langle \underline{b}, \underline{y} \rangle = 0$ für jeden Vektor $\underline{y}$ aus $N(A^H)$. Genauer:

$$\underline{b} \in R(A) \Leftrightarrow \underline{b} \perp N(A^H)$$

- Speziell für Matrizen: Die Gleichung $A\underline{x} = \underline{b}$ hat dann und nur dann eine Lösung, wenn $\underline{b}^H \underline{y} = 0$ für jeden Vektor $\underline{y}$ für den gilt $A^H \underline{y} = 0$.

Orthogonale Unterräume

- **Beweis:** Sei $A\underline{x} = \underline{b}$ und $\underline{v} \in N(A^H)$.
Dann $\langle \underline{b}, \underline{v} \rangle = \langle A\underline{x}, \underline{v} \rangle = \langle \underline{x}, A^H \underline{v} \rangle = \langle \underline{x}, \underline{0} \rangle = 0$

Nehmen wir nun an, dass $\langle \underline{b}, \underline{v} \rangle = 0$ wenn $\underline{v}$ aus $N(A^H)$, aber $A\underline{x} = \underline{b}$ habe keine Lösung. Da $\underline{b}$ nicht aus $R(A)$, nehmen wir an, dass $\underline{b} = \underline{b}_r + \underline{b}_0$, mit $\underline{b}_r$ aus $R(A)$ und $\underline{b}_0$ sei orthogonal zu den Vektoren in $R(A)$. Dann gilt also $\langle A\underline{x}, \underline{b}_0 \rangle = 0$ für alle $\underline{x}$ und somit $A^H \underline{b}_0 = \underline{0}$. Also liegt $\underline{b}_0$ in $N(A^H)$. Und $0 = \langle \underline{b}, \underline{b}_0 \rangle = \langle \underline{b}_r + \underline{b}_0, \underline{b}_0 \rangle = \langle \underline{b}_r, \underline{b}_0 \rangle + \langle \underline{b}_0, \underline{b}_0 \rangle = \langle \underline{b}_0, \underline{b}_0 \rangle \rightarrow \underline{b}_0 = \underline{0}$. $\underline{b}$ ist also aus $R(A)$.

Orthogonale Unterräume

- Damit ist zwar die Existenz bewiesen, nicht aber die Eindeutigkeit der Lösung.
- **Satz 4.6:** Die Lösung von $A\underline{x} = \underline{b}$ ist dann und nur dann eindeutig, wenn die einzige Lösung von $A\underline{x} = \underline{0}$ zu $\underline{x} = \underline{0}$ führt, also $N(A) = \{\underline{0}\}$.

Orthogonale Unterräume

- **Erinnerung:** Die Dimension eines Vektorraumes bezeichnet die Anzahl der linear unabhängigen Vektoren, die nötig ist einen Raum aufzuspannen.
- **Definition 4.9:** Der Rang (engl. rank) einer Matrix A ist durch die Dimension seines Spaltenraums (Zeilenraums) bestimmt.
- **Beispiel 4.9:** Eine $m \times n$ Matrix sei von Rang r :

$$\dim(R(A)) = r; \quad \dim(R(A^H)) = r$$

$$\dim(N(A)) = n - r; \quad \dim(N(A^H)) = m - r$$

Orthogonale Unterräume

- **Beispiel 4.10:**

$$A = \begin{bmatrix} 1 & 4 \\ 5 & 20 \end{bmatrix}; n = 2, m = 2, r = 1$$

$$R(A) = \text{span} \left(\begin{bmatrix} 1 \\ 5 \end{bmatrix} \right); \quad R(A^T) = \text{span} \left(\begin{bmatrix} 1 \\ 4 \end{bmatrix} \right)$$

$$N(A) = \text{span} \left(\begin{bmatrix} -4 \\ 1 \end{bmatrix} \right); \quad N(A^T) = \text{span} \left(\begin{bmatrix} -5 \\ 1 \end{bmatrix} \right)$$

Orthogonale Unterräume

- **Definition 4.10:** Eine $m \times n$ Matrix heißt von vollem Rang, wenn $\text{rang}(A) = \min(m, n)$.
- Ist eine Matrix nicht von vollem Rang, so heißt sie rangdefizient.
- **Satz 4.7:** Für Matrixprodukte AB gilt:

$$N(B) \subset N(AB)$$

$$R(AB) \subset R(A)$$

$$N(A^H) \subset N((AB)^H)$$

$$R((AB)^H) \subset R(B)$$

Orthogonale Unterräume

- **Beweis:** (Nur Teil 1) Wenn $B\underline{x} = \underline{0}$, dann $AB\underline{x} = \underline{0}$. Somit liegt jedes $\underline{x}$ aus $N(B)$ auch in $N(AB)$.
- Man beachte, dass die zweite und vierte Eigenschaft zu folgendem Ergebnis führt:

$$\text{rang}(AB) \leq \text{rang}(A)$$

$$\text{rang}(AB) \leq \text{rang}(B)$$

Projektionen

- Wir hatten bereits herausgefunden, dass das überbestimmte LS Problem $A\underline{x} = \underline{b}$ durch lineare Filterung hervorgeht:

$$\hat{\underline{b}} = A(A^H A)^{-1} A^H \underline{b}$$

$$\underline{e} = (I - A(A^H A)^{-1} A^H) \underline{b}$$

- Nehmen wir an, dass sich $\underline{b}$ durch zwei Anteile darstellen lässt: $\underline{b} = A\underline{y} + \underline{z}$. Dabei liegt der erste Anteil im Spaltenraum $R(A)$ von A und $\underline{z}$ liege im orthogonalen Komplementärraum also im $N(A^H)$.

Projektionen

- Damit gilt für die LS Schätzung:

$$\hat{\underline{b}} = A(A^H A)^{-1} A^H (A\underline{y} + \underline{z}) = A\underline{y}$$

$$\underline{e} = (I - A(A^H A)^{-1} A^H) (A\underline{y} + \underline{z}) = \underline{z}$$

- Dies klärt wieso der Fehler $\underline{e}$ und die LS Schätzung orthogonal aufeinander stehen: sie kommen aus orthogonalen Komplementärräumen.
- Dies bedeutet aber auch, dass die beiden Projektionsmatrizen P und $I - P$ einen Vektor in zwei orthogonale Komplementärräume abbildet (projiziert).

Projektionssatz

- **Satz 4.8:** Sei S ein Hilbert Raum und V ein abgeschlossener Unterraum auf S . Für jeden Vektor $\underline{x}$ aus S existiert eindeutig ein Vektor $\underline{v}_0$ aus V , der $\underline{x}$ am nächsten ist, d.h.,

$$\|\underline{x} - \underline{v}_0\| \leq \|\underline{x} - \underline{v}\|; \text{ für alle } \underline{v} \in V$$

- Darüber hinaus gilt: Dann und nur dann, wenn $\underline{x} - \underline{v}_0$ orthogonal zu V ist, ist $\underline{v}_0$ dem Vektor $\underline{x}$ am nächsten.

Projektionssatz

- **Beweis:** Betrachten wir zunächst die Existenz der Lösung. Sei $\underline{v}_i$ eine Folge von Vektoren in V , die einem Grenzwert zustrebt:

$$\|\underline{x} - \underline{v}_i\| \rightarrow \delta$$

- Hierbei bezeichne δ den folgenden Grenzwert:

$$\delta = \inf_{\underline{v} \in V} \|\underline{x} - \underline{v}\|$$

- Betrachte folgende zwei Vektoren $\underline{v}_i$ und $\underline{v}_j$ aus V :

$$\|\underline{v}_i - \underline{v}_j\|^2 = \|(\underline{v}_i - \underline{x}) + (\underline{x} - \underline{v}_j)\|^2 = 2\|\underline{x} - \underline{v}_j\|^2 + 2\|\underline{x} - \underline{v}_i\|^2 - 4\|\underline{x} - (\underline{v}_j + \underline{v}_i)/2\|^2$$

$$\|(\underline{v}_i - \underline{x}) + (\underline{x} - \underline{v}_j)\|^2 + \|(\underline{v}_i - \underline{x}) - (\underline{x} - \underline{v}_j)\|^2 = 2\|\underline{x} - \underline{v}_j\|^2 + 2\|\underline{x} - \underline{v}_i\|^2$$

Einsetzen, Nachrechnen!!!

Projektionssatz

- Da $\underline{v}_i$ und $\underline{v}_j$ aus V und somit aus S sind, gilt auch, dass $(\underline{v}_i + \underline{v}_j)/2$ in S liegt und somit muss gelten, dass:

$$\|\underline{x} - (\underline{v}_i + \underline{v}_j)/2\| \geq \delta$$

- Damit gilt:

$$\begin{aligned} \|\underline{v}_i - \underline{v}_j\|^2 &= 2\|\underline{x} - \underline{v}_j\|^2 + 2\|\underline{x} - \underline{v}_i\|^2 - 4\|\underline{x} - (\underline{v}_i + \underline{v}_j)/2\|^2 \\ &\leq 2\|\underline{x} - \underline{v}_j\|^2 + 2\|\underline{x} - \underline{v}_i\|^2 - 4\delta^2 \end{aligned}$$

- Da $\underline{v}_i$ aber δ zustrebt, finden wir

$$\|\underline{v}_i - \underline{v}_j\|^2 \leq 2\|\underline{x} - \underline{v}_j\|^2 - 2\delta^2$$

Projektionssatz

- Fordern wir ebenso für die Folge $\underline{v}_j$, dass sie demselben Grenzwert zustrebt, dann folgt daraus, dass

$$\|\underline{v}_i - \underline{v}_j\| \rightarrow 0$$

- Und somit handelt es sich um eine Cauchy Folge.
- Da V ein Hilbertraum ist, existiert der Grenzwert und somit die Lösung $\underline{v}_0$.

Projektionssatz

- Wir wollen zeigen, dass wenn $\underline{v}_0$ die Norm $\|\underline{v}_0 - \underline{x}\|$ minimiert, dann liegt $\underline{v}_0 - \underline{x}$ orthogonal zu V .
- Sei nun $\underline{v}_0$ der zu $\underline{x}$ nächste Vektor.
- Sei $\underline{v}$ ein Vektor der Länge 1. Nehmen wir an, dass keine Orthogonalität gilt, also $\langle \underline{x} - \underline{v}_0, \underline{v} \rangle = \varepsilon > 0$
- Sei $\underline{z} = \underline{v}_0 + \varepsilon \underline{v}$ aus V . Dann gilt:

$$\begin{aligned} \|\underline{x} - \underline{z}\|_2^2 &= \|\underline{x} - \underline{v}_0\|_2^2 - 2 \operatorname{Re} \langle \underline{x} - \underline{v}_0, \varepsilon \underline{v} \rangle + \|\varepsilon \underline{v}\|_2^2 \\ &= \|\underline{x} - \underline{v}_0\|_2^2 - 2|\varepsilon|^2 + |\varepsilon|^2 \|\underline{v}\|_2^2 = \|\underline{x} - \underline{v}_0\|_2^2 - |\varepsilon|^2 \leq \|\underline{x} - \underline{v}_0\|_2^2 \end{aligned}$$
- Dies ist aber ein Widerspruch. Daher muss $\varepsilon = 0$ sein.

Projektionssatz

- Nun muss noch das Gegenteil bewiesen werden. Nehmen wir also an, dass $\underline{v}_0 - \underline{x}$ orthogonal zu V liegt, aber $\underline{v}_0$ nicht der nächste Vektor zu $\underline{x}$ darstellt. Dann muss für jedes $\underline{v}$ aus V aber ungleich $\underline{v}_0$ gelten:

$$\begin{aligned} \|\underline{x} - \underline{v}\|_2^2 &= \|\underline{x} - \underline{v}_0 + \underline{v}_0 - \underline{v}\|_2^2 \\ &= \|\underline{x} - \underline{v}_0\|_2^2 + \|\underline{v}_0 - \underline{v}\|_2^2 \\ &\geq \|\underline{x} - \underline{v}_0\|_2^2 \end{aligned}$$
- Und somit wieder ein Widerspruch.

Projektionssatz

- Eine wichtige Folgerung aus dem Projektionssatz ist, dass jeder Hilbertraum S aus zwei Teilräumen dargestellt werden kann: ein Teilraum V und seinem orthogonalem Komplementärraum $V^\perp$.
- Beachte, dass diese Zerlegung im Hilbertraum wirklich eindeutig ist und abgeschlossen:

$$V \cup V^\perp = S$$

- Eine Eigenschaft, die im allgemeinen bei orthogonalen Komplementärräumen nicht erfüllt ist (wenn die Eigenschaft der Abgeschlossenheit fehlt).

Faktorisierung

- Ein Gleichungssystem der Form $A\underline{x} = \underline{b}$ kann auf vielerlei Art gelöst werden. Dabei ist die numerische Genauigkeit oft wesentlich für die Qualität des Ergebnisses.
- Es existieren zahlreiche Methoden, zur Matrixzerlegung, die das allgemeine Problem $A\underline{x} = \underline{b}$ in ein äquivalentes Problem $B\underline{x} = \underline{c}$ umwandeln, bei dem die Matrix B besondere Eigenschaften hat, so dass das System sehr leicht zu lösen ist.
- Im Folgenden werden wir nur die wesentlichen Ideen vorstellen, nicht aber auf die Details der Verfahren eingehen.

Faktorisierung

- **LU:** steht für „lower-“ und „upper-triangular“. $A=LU$ lässt sich einfacher lösen, da $LU\mathbf{x}=\mathbf{b}$: $U\mathbf{x}=\mathbf{c}$ und $L\mathbf{c}=\mathbf{b}$, bedeutet, zwei Gleichungssysteme, die einfach zu lösen sind.
- **Cholesky:** Eine Spezialform der LU Zerlegung besteht für hermitesche, positiv-definite Matrizen $A=LL^H$.
Allgemein lassen sich Matrizen in die Formen LDL^H , UDU^H oder QDQ^H zerlegen. Hierbei ist Q eine unitäre Matrix: $QQ^H=I$ und D eine Diagonalmatrix. Im Falle von QDQ^H spricht man von Eigenwertzerlegung.
- **QR:** $A=QR$. Hierbei ist Q eine unitäre Matrix: $QQ^H=I$ und R ist eine obere Dreiecksmatrix. $A=LU$ lässt sich einfacher lösen, da $QR\mathbf{x}=\mathbf{b}$: $R\mathbf{x}=\mathbf{c}$ und $Q\mathbf{c}=\mathbf{b} \rightarrow \mathbf{c}=Q^H\mathbf{b}$, bedeutet, zwei Gleichungssysteme, die einfach zu lösen sind.

Faktorisierung

- **SVD:** Singular Value Decomposition. $A=U\Sigma V$ mit den beiden unitären Matrizen U und V und der Diagonalmatrix Σ .
- Im Folgenden werden wir die Eigenwertzerlegung vertiefen und dann die Singulärwertzerlegung zeigen.

Eigenwertzerlegung

- Sei A ein $m \times m$ Matrix in $\mathbb{C}$. Betrachte die Gleichung

$$A\mathbf{x}=\lambda\mathbf{x}$$
oder äquivalent hierzu

$$(A-\lambda I)\mathbf{x}=\mathbf{0}.$$
- Hierbei ist nicht die triviale Lösung $\mathbf{x}=\mathbf{0}$ gesucht sondern der Nullraum von $(A-\lambda I)$.
- Die speziellen Werte λ , die einen nichttrivialen Nullraum erzeugen, heißen Eigenwerte. Die zugehörigen Vektoren $\mathbf{x}$, Eigenvektoren.

Eigenwertzerlegung

- **Definition 4.11:** Das Polynom in λ , das durch die Determinante von $(A-\lambda I)$ erzeugt wird, heißt charakteristisches Polynom.
Die Determinante $\det(A-\lambda I)=0$ heißt charakteristische Gleichung von A .
Die Wurzeln der charakteristischen Gleichung heißen Eigenwerte. Die Menge aller Eigenwerte heißt Spektrum von A .

Eigenwertzerlegung

- **Beispiel 4.11:** Ein lineares, zeitinvariantes System werde durch folgende Gleichung im Zustandsraum beschrieben:

$$\begin{aligned}\underline{x}_{k+1} &= A\underline{x}_k + B\underline{u}_k \\ \underline{y}_k &= C\underline{x}_k \\ &= H(z)\underline{u}_k \\ &= C(zI - A)^{-1}B\underline{u}_k\end{aligned}$$

- Durch die Matrixinversion von $(zI - A)$ bestimmt die Determinante $\det(zI - A)$ das Verhalten des Systems.

Eigenwertzerlegung

- **Lemma 4.2:** Wenn die Eigenwerte einer Matrix A alle unterschiedlich sind, so sind die zugehörigen Eigenvektoren linear unabhängig.

- **Beweis:** Wir beginnen mit $m=2$ und dem Gegenteil: Die Eigenvektoren $\underline{x}_1$ und $\underline{x}_2$ seien linear abhängig also

$$\begin{aligned}c_1 \underline{x}_1 + c_2 \underline{x}_2 &= \underline{0} \\ c_1 A \underline{x}_1 + c_2 A \underline{x}_2 &= c_1 \lambda_1 \underline{x}_1 + c_2 \lambda_2 \underline{x}_2 \\ - | & \quad c_1 \lambda_2 \underline{x}_1 + c_2 \lambda_2 \underline{x}_2 = \underline{0} \\ c_1 (\lambda_1 - \lambda_2) \underline{x}_1 &= \underline{0}\end{aligned}$$

- Da λ_1 und λ_2 verschieden und $\underline{x}_1$ ungleich dem Nullvektor, muss $c_1 = 0$ sein.

Eigenwertzerlegung

- Ähnliche Argumentation liefert $c_2 = 0$. Das beweist aber, dass die beiden Vektoren linear unabhängig sein müssen.
- Für $m > 2$ betrachtet man immer den Fall, dass genau zwei Vektoren linear abhängig sind und beweist den Widerspruch entsprechend.
- Wenn die Eigenwerte nicht unterschiedlich sind, dann kann es linear unabhängige Eigenvektoren geben oder nicht. Betrachte hierzu die Matrizen:

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}; \quad B = \begin{bmatrix} 4 & 1 \\ 0 & 4 \end{bmatrix};$$

Eigenwertzerlegung

- Betrachte die Zerlegung $A = SDS^{-1}$, mit der Diagonalmatrix D .
- Nehmen wir zunächst an, dass A m linear unabhängige Eigenvektoren habe.
- Damit gilt:

$$[A\underline{x}_1, A\underline{x}_2, \dots, A\underline{x}_m] = [\lambda_1 \underline{x}_1, \lambda_2 \underline{x}_2, \dots, \lambda_m \underline{x}_m] = AS = SD.$$

- Sind die Eigenvektoren linear unabhängig, so kann S invertiert werden, und wir finden:

$$A = SDS^{-1}$$

Eigenwertzerlegung

- **Hinweis:** Solche Matrixtransformationen bei denen die Eigenwerte erhalten bleiben nennt man Ähnlichkeitstransformationen. Zwei Matrizen heißen ähnlich, wenn sie die selben Eigenwerte haben.

- Vorteil der Transformation:

$$A^n = S D^n S^{-1}$$

$$f(A) = \sum_i f_i A^i = S \left(\sum_i f_i D^i \right) S^{-1}$$

$$e^A = S \left(\sum_{i=0}^{\infty} \frac{D^i}{i!} \right) S^{-1} = S e^D S^{-1}$$

Jordanform

- Bei Eigenvektoren, die nicht linear unabhängig sind, ist eine Diagonalisierung nicht möglich! Allerdings ist eine fast-Diagonalform möglich, die so genannte Jordan Form:

$$A = T J T^{-1}$$

- Die Matrix J ist von Blockdiagonalform, mit den Blöcken J_i :

$$J_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}$$

Jordanform

- **Beispiel 4.12:** Betrachte die folgende Matrix:

$$B = \begin{pmatrix} 3 & 0 & 1 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

- Sie hat einen einzigen Eigenwert $\lambda=3$, und zwei linear unabhängige Eigenvektoren:

$$\underline{x}_1 = [1, 0, 0]^T; \underline{x}_2 = [0, 1, 0]^T;$$

Jordanform

- **Noch Beispiel 4.12:** Damit lautet die Jordanform der Matrix B:

$$J(B) = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

- Es gilt: $B = S J S^{-1}$ mit: $S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}$
- Es gilt: $B^n = S J^n S^{-1}$
- Aber J^n ist nicht diagonal oder Jordanförmig!

Minimalpolynom

- **Satz 4.9:** (Cayley Hamilton): Jede quadratische Matrix A erfüllt ihre charakteristische Gleichung.
- **Definition 4.12:** Ein Polynom heißt anhilierendes Polynom einer quadratischen Matrix A , wenn gilt: $f(A)=0$.
- **Definition 4.13:** Dasjenige anhilierende (monische) Polynom von A von kleinstem Grad heißt Minimalpolynom von A .

Minimalpolynom

- **Beispiel 4.13:** Betrachte folgende Matrizen

$$A_1 = \begin{bmatrix} 4 & & \\ & 4 & \\ & & 4 \end{bmatrix}; \quad A_2 = \begin{bmatrix} 5 & 1 & \\ & 5 & \\ & & 5 \end{bmatrix}; \quad A_3 = \begin{bmatrix} 6 & 1 & \\ & 6 & 1 \\ & & 6 \end{bmatrix};$$

- Die Minimalpolynome dazu lauten:

$$f_1(x) = x - 4; \quad f_2(x) = (x - 5)^2; \quad f_3(x) = (x - 6)^3;$$

- Wir erkennen den Zusammenhang mit der Größe der Jordanblöcke.

Hermiteische Matrizen

- Gilt $A=A^T$, für A aus $\mathbb{R}$, so heißt die Matrix symmetrisch.
- Gilt $A=A^H$, für A aus $\mathbb{C}$, so heißt die Matrix hermitesch.
- Solche Matrizen tauchen üblicherweise in Form von Kovarianzmatrizen $R_{xx}=E[\mathbf{x}\mathbf{x}^H]$ auf.
- **Lemma 4.3:** Die Eigenwerte von hermiteschen Matrizen sind reell-wertig.
- **Lemma 4.4:** Die Eigenvektoren zu unterschiedlichen Eigenwerten hermitescher Matrizen sind orthogonal.

Hermiteische Matrizen

- **Beweis:** Seien λ_1 und λ_2 zwei unterschiedliche Eigenwerte mit korrespondierenden Eigenvektoren $\underline{x}_1$ und $\underline{x}_2$. Dann gilt:
 $\langle A\underline{x}_1, \underline{x}_2 \rangle = \langle \underline{x}_1, A\underline{x}_2 \rangle = \langle \underline{x}_1, \lambda_2 \underline{x}_2 \rangle = \lambda_2 \langle \underline{x}_1, \underline{x}_2 \rangle = \lambda_1 \langle \underline{x}_1, \underline{x}_2 \rangle$
 also: $(\lambda_2 - \lambda_1) \langle \underline{x}_1, \underline{x}_2 \rangle = 0$ und somit $\langle \underline{x}_1, \underline{x}_2 \rangle = 0$.
- Beachte: Jede hermitesche $m \times m$ Matrix A kann durch eine unitäre Matrix diagonalisiert werden. Die unitäre Matrix besteht einfach aus den orthonormalen Eigenvektoren.

Unterraumtechniken

- **Definition 4.14:** Sei A eine $m \times m$ Matrix und S eine Teilmenge von $R(A)$. S heißt invarianter Unterraum von A , wenn es für jedes $\underline{x}$ aus S ein $A\underline{x}$ aus S gibt.
- **Beispiel 4.14:** Eine $m \times m$ Matrix A habe k unterschiedliche Eigenwerte mit den zugehörigen Eigenvektoren $\underline{x}_i$, $i=1..m$. Sei $X=[\underline{x}_1, \underline{x}_2, \dots, \underline{x}_m]$ und X_i eine Teilmenge der Eigenvektoren, die zu dem Eigenwert λ_i gehören. $\text{Span}(X_i)$ bezeichnet invariante Unterräume von A .

Unterraumtechniken

- **Satz 4.10:** Sei A eine $m \times m$ hermitesche Matrix mit k unterschiedlichen Eigenwerten. Dann gilt:

$$\text{Spektrale Dekomposition : } A = \sum_{i=1}^k \lambda_i P_i$$

$$\text{Identität : } I = \sum_{i=1}^k P_i$$

$$\text{mit : } P_i = \sum_{\underline{x}_j \in X_i} \underline{x}_j \underline{x}_j^H$$

- Die Matrizen P_i sind Projektionsmatrizen auf den Raum $\text{span}(X_i)$.

Unterraumtechniken

- Man betrachte die Wirkung der Matrix A auf einen beliebigen Vektor $\underline{x}$:

$$A\underline{x} = \sum_{i=1}^k \lambda_i P_i \underline{x} = \sum_{i=1}^k \underbrace{\lambda_i}_{\text{Streckung der Komponenten}} \underbrace{\sum_{\underline{x}_j \in X_i} \underline{x}_j \underline{x}_j^H \underline{x}}_{\text{Projektion von } \underline{x} \text{ auf die einzelnen Komponenten des Unterraums}}$$

Unterraumtechniken

- **Beispiel 4.15:** Betrachte die Matrix $A=[8.2, -2.4; -2.4, 6.8]$ mit den beiden Eigenwerten $\lambda_1=5$ und $\lambda_2=10$ mit den zugehörigen orthogonalen Eigenvektoren

$$\underline{x}_1 = \begin{bmatrix} 3 \\ 4 \end{bmatrix}; \quad \underline{x}_2 = \begin{bmatrix} -4 \\ 3 \end{bmatrix}$$

$$P_1 = \frac{1}{25} \begin{bmatrix} 3 \\ 4 \end{bmatrix} \begin{bmatrix} 3 & 4 \end{bmatrix}^T = \frac{1}{25} \begin{bmatrix} 9 & 12 \\ 12 & 16 \end{bmatrix}$$

$$P_2 = \frac{1}{25} \begin{bmatrix} -4 \\ 3 \end{bmatrix} \begin{bmatrix} -4 & 3 \end{bmatrix}^T = \frac{1}{25} \begin{bmatrix} 16 & -12 \\ -12 & 9 \end{bmatrix}$$

$$A = 5P_1 + 10P_2$$

Unterraumtechniken

- Man betrachte den Ausdruck

$$\underline{x}^H A \underline{x} = \sum_{i=1}^k \lambda_i \underline{x}^H P_i \underline{x} = \sum_{i=1}^k \lambda_i \sum_{\underline{x}_j \in X_j} \underline{x}^H \underline{x}_j \underline{x}_j^H \underline{x}$$

- Wählt man verschiedene Eigenvektoren $\underline{x} = \underline{x}_j$ aus, so erhält man gerade die Eigenwerte; beim größten Eigenvektor erhält man also den größten Eigenwert, usw.

$$\max_{\underline{x}} \frac{\underline{x}^H A \underline{x}}{\underline{x}^H \underline{x}} = \lambda_{\max}; \quad \min_{\underline{x}} \frac{\underline{x}^H A \underline{x}}{\underline{x}^H \underline{x}} = \lambda_{\min}$$

Unterraumtechniken

- Allgemein heißt der Ausdruck

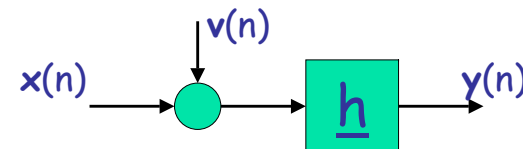
$$r(\underline{x}) = \frac{\underline{x}^H A \underline{x}}{\underline{x}^H \underline{x}}$$

- Rayleigh Quotient. Beachte, dass dieser Ausdruck nur Sinn bei hermiteschen Matrizen macht.

Eigenfilter

- Bekannt ist das Signalangepasste Filter, welches das Signal-Rausch Verhältnis für ein determiniertes Signal maximiert.
- Betrachtet man hingegen das maximale Signal-Rausch-Verhältnis bei Zufallssignalen, so spricht man von Eigenfilter.

Eigenfilter



- Für ein Zufallssignal $\underline{x}(n)$ gilt:

- Empfangsleistung $P = \underline{h}^H E[\underline{x}^H \underline{x}] \underline{h} = \underline{h}^H R_{xx} \underline{h}$
- Rauschleistung: $N = \underline{h}^H E[\underline{v}^H \underline{v}] \underline{h} = \sigma_v^2 \underline{h}^H \underline{h}$
- Maximiere das Signal Rauschverhältnis

$$\max_{\underline{h}} \frac{P}{N} = \max_{\underline{h}} \frac{\underline{h}^H R_{xx} \underline{h}}{\sigma_v^2 \underline{h}^H \underline{h}} = \frac{\lambda_{\max}(R_{xx})}{\sigma_v^2}$$

- Die optimale Lösung ergibt sich aus dem Eigenvektor zu $\lambda_{\max}$.

Filterbeispiel

- Die eben erläuterte Technik kann zum Filterentwurf genutzt werden. Sei ein linear phasiges Filter mit $2N+1$ Koeffizienten zu bestimmen, für das ein Amplitudengang $H_d(e^{j\Omega})$ vorgegeben ist.

$$H(e^{j\Omega}) = e^{-jN\Omega} H_R(e^{j\Omega}) = e^{-jN\Omega} \sum_{n=0}^N b_n \cos(n\Omega) = e^{-jN\Omega} \underline{b}^T \underline{c}(\Omega)$$

- Sei ein Tiefpassfilter mit den Grenzfrequenzen ω_p, ω_s zu entwerfen, also mit Amplitudengang

$$H_d(e^{j\Omega}) = \begin{cases} 1 & ; 0 \leq \Omega \leq \Omega_p \\ 0 & ; \Omega_s \leq \Omega \leq \pi \end{cases}$$

Filterbeispiel

- Im Stoppband gilt:

$$E_s = \int_{\Omega_s}^{\pi} |H_r(e^{j\Omega}) - H_d(e^{j\Omega})|^2 d\Omega$$

$$= \frac{1}{\pi} \underline{b}^T \int_{\Omega_s}^{\pi} \underline{c}(\Omega) \underline{c}^T(\Omega) d\Omega \underline{b} = \underline{b}^T P \underline{b}$$

- Mit der Matrix P:

$$P_{ij} = \frac{1}{\pi} \int_{\Omega_s}^{\pi} \cos(i\Omega) \cos(j\Omega) d\Omega$$

Filterbeispiel

- Im Passband gilt $H_d(e^{j\Omega})=1$, oder äquivalent $\underline{b}^T \underline{1}=1$. Die Abweichung hierzu lautet:
 $1 - \underline{b}^T \underline{c}(\Omega) = \underline{b}^T [1 - \underline{c}(\Omega)]$.

- Die darin enthaltene (Fehler-) Energie sollte minimal werden, also:

$$E_p = \frac{1}{\pi} \underline{b}^T \int_{\Omega_p}^{\pi} [1 - \underline{c}(\Omega)] [1 - \underline{c}(\Omega)]^T d\Omega \underline{b} = \underline{b}^T Q \underline{b}$$

- Das gesamte Filterproblem lässt sich also darstellen als:

$$J(\alpha) = \alpha E_s + (1-\alpha) E_p = \underline{b}^T R \underline{b} = \alpha \underline{b}^T P \underline{b} + (1-\alpha) \underline{b}^T Q \underline{b}$$

Filterbeispiel

- Offensichtlich haben wir noch Freiraum in der Wahl von $\underline{b}$. Diesen können wir beispielsweise durch eine Normierung der Form $\underline{b}^T \underline{b}=1$ einschränken.

- Das Filterproblem lautet dann:

Minimiere $\underline{b}^T R \underline{b}$ mit Nebenbedingung $\underline{b}^T \underline{b}=1$.

- Oder, äquivalent:

$$\min_{\underline{b}} \frac{\underline{b}^T R \underline{b}}{\underline{b}^T \underline{b}}$$

Weitere Unterraumtechniken

- Viele moderne Verfahren der digitalen Signalverarbeitung beruhen auf Unterraumtechniken. Die wichtigsten sind:
 - Pisarenkos Harmonische Dekomposition (PHD)
 - MUSIC
 - ESPRIT
- Sie werden im Folgenden vorgestellt.

Weitere Unterraumtechniken

- Betrachte folgendes Signalmodell:

$$\mathbf{x}(t) = \sum_{i=1}^p a_i e^{j2\pi f_i t} + \mathbf{w}(t)$$

unbekannte
Amplituden
aus $\mathbf{C}$

unbekannte
Frequenzen

weißes Rauschen
unbekanntes σ_w^2

Weitere Unterraumtechniken

- Durch das additive Rauschen wird das Signal $\mathbf{x}(k)$ ein Zufallsprozess.
- Wir tasten das Signal an M ($p < M$) Stellen ab und erhalten:

$$\underline{s}_i = [1, e^{j2\pi f_i T}, e^{j2\pi f_i 2T}, \dots, e^{j2\pi f_i (M-1)T}]^T$$

$$\underline{R}_{\mathbf{xx}} = E[\underline{\mathbf{x}}_k \underline{\mathbf{x}}_k^H] = \sum_{i=1}^p |a_i|^2 \underline{s}_i \underline{s}_i^H + \underline{R}_{\mathbf{ww}}$$

$$= \mathbf{S} \mathbf{A} \mathbf{S}^H + \underline{R}_{\mathbf{ww}}$$

Weitere Unterraumtechniken

- Beachte, dass $\mathbf{S}$ eine Vandermonde Matrix ist.
- Der Vektorraum $\text{span}(\underline{s}_1, \dots, \underline{s}_p)$, der durch $\mathbf{S}$ aufgespannt wird, stellt einen Unterraum des Signals $\underline{\mathbf{x}}_k$ dar. Er heißt Signal-Unterraum.
- Setzt man das Rauschen zu Null, so kann man nur $p < M$ von Null verschiedene Eigenwerte finden mit den zugehörigen Eigenvektoren (Haupteigenvektoren):

$$\underline{R}_{\mathbf{xx}}|_{\mathbf{w}(t)=0} = \sum_{i=1}^p \lambda_i \underline{u}_i \underline{u}_i^H$$

$$\text{span}\{\underline{s}_1, \underline{s}_2, \dots, \underline{s}_p\} = \text{span}\{\underline{u}_1, \underline{u}_2, \dots, \underline{u}_p\}$$

Weitere Unterraumtechniken

- Dies gibt also prinzipiell die Möglichkeit ohne Kenntnis der einzelnen Frequenzen f_i den Unterraum des Signals zu bestimmen.
- Mit seiner Kenntnis lassen sich dann in einem zweiten Schritt, die Frequenzen bestimmen, bzw. die Signalqualität verbessern.
- Nehmen wir weiterhin weißes Rauschen $\mathbf{w}(t)$ an:

$$\mathbf{R}_{\mathbf{xx}} = \mathbf{S}\mathbf{P}\mathbf{S}^H + \sigma_w^2 \mathbf{I}$$

Weitere Unterraumtechniken

- Wir erkennen nun, dass bei Vorhandensein von Rauschen die ersten p Eigenwerte um σ_w^2 verschoben sind. Die zugehörigen Eigenvektoren bleiben aber unverändert.
- Hinzu kommen nun noch $M-p$ Eigenwerte der Größe σ_w^2 mit zugehörigen Eigenvektoren $\underline{u}_{p+1}, \dots, \underline{u}_M$. Diese sind allein durch das Rauschen bestimmt.
- Man nennt

$$\mathbf{N} = \text{span}\{\underline{u}_{p+1}, \underline{u}_{p+2}, \dots, \underline{u}_M\}$$

- den Rausch-Unterraum.
- Beachte, dass jeder Vektor des Signal-Unterraums orthogonal zu jedem Vektor des Rausch-Unterraums ist.

Pisarenkos Harmonische Dekomposition

- Zunächst werden aus der AKF Matrix $\mathbf{R}_{\mathbf{xx}}$ die Eigenwerte und Eigenvektoren bestimmt. Dadurch kann p und σ_w^2 bestimmt werden.
- Da die Eigenvektoren aus dem Rausch-Unterraum orthogonal zum Signal Unterraum sind, kann man folgern: $\underline{s}_i^H \underline{u}_M = 0$; für $i=1,2,\dots,p$
- Dadurch erhält man ein Bestimmungspolynom für die unbekannten Frequenzen.
- Mit diesen können schließlich die Leistungen p_i bestimmt werden.

Pisarenkos Harmonische Dekomposition

- Mit diesen können schließlich die Leistungen p_i bestimmt werden.

$$\begin{bmatrix} e^{j2\pi f_1} & e^{j2\pi f_2} & e^{j2\pi f_3} & \dots & e^{j2\pi f_p} \\ e^{j2\pi 2f_1} & e^{j2\pi 2f_2} & e^{j2\pi 2f_3} & \dots & e^{j2\pi 2f_p} \\ e^{j2\pi 3f_1} & e^{j2\pi 3f_2} & e^{j2\pi 3f_3} & \dots & e^{j2\pi 3f_p} \\ \vdots & & & \ddots & \vdots \\ e^{j2\pi pf_1} & \dots & e^{j2\pi pf_3} & \dots & e^{j2\pi pf_p} \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ \vdots \\ p_p \end{bmatrix} = \begin{bmatrix} r_{\mathbf{xx}}(1) \\ r_{\mathbf{xx}}(2) \\ r_{\mathbf{xx}}(3) \\ \vdots \\ r_{\mathbf{xx}}(p) \end{bmatrix}$$

- Problematisch bei dieser Methode ist in erster Linie die ungenaue Bestimmung der Frequenzen.

MUSIC

- MUSIC=MUltiple SIgnal Classification.
- Betrachte den Vektor:

$$\underline{s}^T(f) = [1, e^{j2\pi f}, e^{j2\pi 2f}, \dots, e^{j2\pi (M-1)f}]$$

$$\underline{s}^H(f) \underline{u}_l = 0; l = p+1, \dots, M \text{ und } f = f_i$$

- Berechnet man den Ausdruck

$$P(f) = \frac{1}{\sum_{k=p+1}^M |\underline{s}^H(f) \underline{u}_k|^2}$$

- So wird dieser für die gesuchten Frequenzen maximal. Danach kann man wie bei PHD verfahren.

ESPRIT

- ESPRIT=Estimation of Signal Parameters via Rotational Invariance Techniques.
- Betrachtet werden die beiden Kovarianzmatrizen:

$$R = E[\underline{x}_k \underline{x}_k^H] \quad Q = E[\underline{x}_k \underline{x}_{k+1}^H]$$

$$R = SPS^H + \sigma_w^2 I; \quad Q = SP\Phi S^H + \sigma_w^2 E$$

$$\Phi = \text{diag}(e^{j2\pi f_1}, e^{j2\pi f_2}, \dots, e^{j2\pi f_p})$$

$$E = \begin{bmatrix} 0 & & & & \\ 1 & 0 & & & \\ & \ddots & \ddots & & \\ & & & 1 & 0 \end{bmatrix}$$

ESPRIT

- Über die Eigenwertzerlegung von R wird wie bisher p und σ_w^2 bestimmt.
- Dadurch erhält man SPS^H sowie $SP\Phi S^H$.
- Betrachte nun $SP[I - \lambda\Phi]S^H \underline{u} = \underline{0}$.
- Die verallgemeinerten Eigenwerte dieser Gleichung sind die gesuchten Werte $e^{j2\pi f_i}$.
- Verallgemeinerte Eigenwerte werden durch die folgende Gleichung angegeben: $SPS^H \underline{u} = \lambda SP\Phi S^H \underline{u}$
- Tip: versuche Matlab $D = \text{eig}(A, B)$.

ESPRIT

- Anwendungen der Verfahren
 - Beachte, dass $h(t-\tau) \rightarrow H(e^{j\Omega\tau}) e^{j\Omega\tau}$
 - Damit sind alle Bestimmungen von zeitl. Verzögerungen auf Frequenzbestimmungen zurückgeführt, bsp: Radartechniken
 - AoA (Angle of Arrival) und AoD (Angle of Departure) Bestimmungen bei Funkmessungen

Singulärwertzerlegung

- **Satz 4.11:** Jede Matrix A aus $\mathbb{C}^{m \times n}$ kann in folgende Form zerlegt werden: $A = U \Sigma V^H$, mit den unitären Matrizen U aus $\mathbb{C}^{m \times m}$ und V aus $\mathbb{C}^{n \times n}$ sowie der Diagonalmatrix Σ aus $\mathbb{R}^{m \times n}$, mit $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_p)$, $p = \min(m, n)$.
- Diese spezielle Faktorisierung heißt Singulärwertzerlegung (engl.: Singular Value Decomposition = SVD). Die Diagonalwerte der Matrix Σ heißen Singulärwerte.

Singulärwertzerlegung

- **Beweis:** Betrachte die Eigenwertzerlegung einer Matrix $A^H A$: $A^H A V = V D$
- Die Eigenwerte in D seien so geordnet, dass $\lambda_1, \dots, \lambda_r > 0$ und $\lambda_{r+1} = \dots = \lambda_p = 0$.
- Damit lassen sich die folgenden r Vektoren bilden:

$$\underline{u}_i = \frac{A \underline{v}_i}{\sqrt{\lambda_i}}$$

- Wir finden, dass $\langle \underline{u}_i, \underline{u}_j \rangle = \delta_{i-j}$.

Singulärwertzerlegung

- Die Menge $\{\underline{u}_1, \dots, \underline{u}_r\}$ kann orthogonal erweitert werden (beispielsweise durch die Gram Schmidt Methode). Damit erhält man mit $U = [\underline{u}_1, \dots, \underline{u}_p]$: $U^H U = I$.
- Offensichtlich sind die Vektoren in U die Eigenvektoren zu $A A^H$. Dies ist klar für die von Null verschiedenen Eigenwerte. Für die nullwertigen Eigenwerte müssen die Eigenvektoren aus dem Nullraum von $A A^H$ kommen. Dies ist durch die Konstruktion gegeben.

Singulärwertzerlegung

- Damit gilt aber für die Elemente von $U A V^H$:

$$i = 1..r \quad : \underline{u}_i^H A \underline{v}_j = \frac{1}{\sqrt{\lambda_i}} \underline{v}_i^H A^H A \underline{v}_j = \sqrt{\lambda_i} \delta_{i-j}$$

$$i = r+1..p : A A^H \underline{u}_i = \underline{0}$$

- Somit liegt $A^H \underline{u}_i$ im Nullraum von A und im Spaltenraum von A^H . Für $A^H \underline{u}_i = \underline{0}$ gilt aber auch $\underline{v}_j^H A^H \underline{u}_i = \underline{u}_j^H A \underline{v}_i = 0$. Also ist $U^H A V = D$ diagonal mit den von Null verschiedenen Elementen $\sqrt{\lambda_j}$, $j = 1..r$.

Singulärwertzerlegung

- Beachte: $m=2, n=3: \Sigma = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \end{bmatrix}$
 $m=3, n=2: \Sigma = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \end{bmatrix}$

- Sei $B=A^H A$. Dann gilt: $BV=V D_1 \rightarrow n \times n$
 $=V \Sigma^T U^H U \Sigma V^H V$
 $=V \Sigma^T \Sigma$
- Ebenso gilt für $B=AA^H: BU=UD_2=U \Sigma \Sigma^T \rightarrow m \times m$

Singulärwertzerlegung

- Beachte weiterhin, dass wenn A aus $\mathbb{R}^{m \times n}$ ist, dann sind alle Matrizen aus $\mathbb{R}$.
- Spektralzerlegung:

$$A = \sum_{i=1}^p \sigma_i \underline{u}_i \underline{v}_i^H = \sum_{i=1}^r \sigma_i \underline{u}_i \underline{v}_i^H$$

- Matrix Norm:
 - Frobenius Norm: $\|A\|_F^2 = \sigma_1^2 + \sigma_2^2 + \dots + \sigma_p^2$.
 - l_2 Norm: $\|A\|_2^2 = \sigma_{\max}^2$

Singulärwertzerlegung

- Sei A hermitesch. Dann gilt: $A=UDU^T=U\Sigma U^T$.
- In diesem Fall sind die Eigenwerte und die Singulärwerte identisch.
- Der Rang einer Matrix ist gegeben durch die r von Null verschiedenen Singulärwerte: $\text{Rang}(A)=r$.
- Partitioniere $U=[U_1 \ U_2]$

$$\begin{aligned} R(A) &= \{ \underline{b} \in C^m : \underline{b} = A \underline{x} \} \\ &= \{ \underline{b} \in C^m : \underline{b} = U \Sigma V^H \underline{x} \} \\ &= \{ \underline{b} \in C^m : \underline{b} = U \Sigma \underline{y} \} \\ &= \{ \underline{b} \in C^m : \underline{b} = U_1 \underline{y} \} = \text{span}(U_1) \end{aligned}$$

Singulärwertzerlegung

- Damit finden wir eine neue Erklärung der vier fundamentalen Unterräume:

$$R(A) = \text{span}(U_1)$$

$$N(A) = \text{span}(V_2)$$

$$R(A^H) = \text{span}(V_1)$$

$$N(A^H) = \text{span}(U_2)$$

Singulärwertzerlegung

- Betrachte noch einmal LS Problem

$$\begin{aligned}\min \|A\underline{x} - \underline{b}\|_2^2 &= \min \|U \Sigma V^H \underline{x} - \underline{b}\|_2^2 \\ &= \min \|\Sigma V^H \underline{x} - U^H \underline{b}\|_2^2 \\ &= \min \|\Sigma \tilde{\underline{x}} - \tilde{\underline{b}}\|_2^2\end{aligned}$$

- Durch die besondere Struktur von Σ , werden die unteren $r+1..p$ Zeilen ausgeblendet.

$$\Sigma = \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & 0 \end{bmatrix}; \quad \Sigma^\# = \begin{bmatrix} 1/\sigma_1 & & & \\ & \ddots & & \\ & & 1/\sigma_r & \\ & & & 0 \end{bmatrix};$$

Singulärwertzerlegung

- Also:

$$\begin{aligned}\tilde{\underline{x}} &= \Sigma^\# \tilde{\underline{b}} \\ V^H \underline{x} &= \Sigma^\# U^H \underline{b} \\ \underline{x} &= V \Sigma^\# U^H \underline{b} = (A^H A)^{-1} A^H \underline{b}\end{aligned}$$

- Damit verstehen wir auch wieso die LS Lösung eine Minimum Norm Lösung ist.

Singulärwertzerlegung

- Anwendungen der SVD:
 - Unterraumtechniken wie PHD, MUSIC und ESPRIT.
 - Total Least Squares: $\|A\underline{x} - \underline{b}\|$ mit gestörter Beobachtungs-Matrix A .
 - Lösung numerisch empfindlicher Probleme.

Konditionszahlen

- Betrachte ein Matrixproblem, bei dem einzelne Singulärwerte sehr klein sind:

$$\Sigma = \begin{bmatrix} \sigma_{\max} & & & \\ & \ddots & & \\ & & \sigma_l & \\ & & & \varepsilon \end{bmatrix}; \Sigma^{-1} \neq \Sigma^\#$$

- Bei der Inversion werden gerade diesen kleinen Werte enorm verstärkt und können zu numerischen Fehlern führen. Besser ist es, sie zu unterdrücken und die Pseudoinverse zu verwenden.

Konditionszahlen

- Beachte, dass durch diese Unterdrückung der kleinsten Eigenwerte auch der Rang der Matrix reduziert wird.
- Diese Methoden heißen daher Rangreduzierende Methoden.
- Ein Maß, dass die Qualität einer Matrix in Hinblick auf ihre Invertierbarkeit angibt, ist die Konditionszahl:

$$\kappa(A) = \|A\|_2 \|A^{-1}\|_2$$

Konditionszahlen

- Die kleinstmögliche Konditionszahl ist Eins! In diesem Fall handelt es sich um eine Matrix proportional zur Identitätsmatrix.
- Ansonsten gilt $\kappa > 1$.
- Da für reguläre Matrizen A gilt:

$$\sigma_{\max} = \max_{\|x\|_2=1} \frac{\|Ax\|_2}{\|x\|_2}$$

$$1/\sigma_{\min} = \max_{\|x\|_2=1} \frac{\|A^{-1}x\|_2}{\|x\|_2}$$

$$= 1/\min_{\|x\|_2=1} \frac{\|Ax\|_2}{\|x\|_2}$$

$$\kappa(A) = \frac{\sigma_{\max}}{\sigma_{\min}}$$

Konditionszahlen

- Für hermitesche Matrizen A ist die Konditionszahl auch durch die Eigenwerte gegeben:

$$\kappa(A) = \frac{\lambda_{\max}}{\lambda_{\min}}$$

- Ansonsten kann man auch alternativ die Wurzeln der Eigenwerte von $A^H A$ berechnen:

$$\kappa(A) = \sqrt{\frac{\lambda_{\max}(A^H A)}{\lambda_{\min}(A^H A)}}$$

Konditionszahlen

- Betrachte folgendes, gestörtes Problem:

$$A(\underline{x} + \Delta \underline{x}) = \underline{b} + \Delta \underline{b}$$

- Man findet, dass

$$\frac{\|\Delta \underline{x}\|}{\|\underline{x}\|} \leq \kappa(A) \frac{\|\Delta \underline{b}\|}{\|\underline{b}\|}$$

- Die Konditionszahl gibt also an, wie sehr sich ein Fehler auf der einen Seite der Gleichung auf die andere Seite der Gleichung auswirkt.