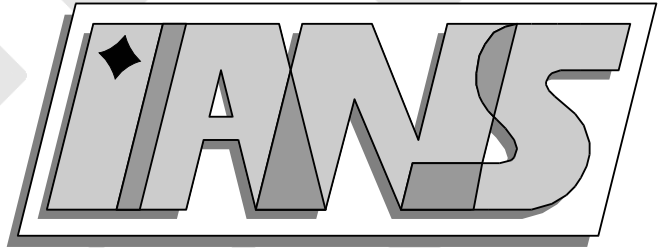


**Universität
Stuttgart**



Reduzierte-Basis-Methoden

Bernard Haasdonk

**Berichte aus dem Institut für
Angewandte Analysis und Numerische Simulation**

Preprint 2011/004

Universität Stuttgart

Reduzierte-Basis-Methoden

Bernard Haasdonk

**Berichte aus dem Institut für
Angewandte Analysis und Numerische Simulation**

Preprint 2011/004

Institut für Angewandte Analysis und Numerische Simulation (IANS)
Fakultät Mathematik und Physik
Fachbereich Mathematik
Pfaffenwaldring 57
D-70 569 Stuttgart

E-Mail: ians-preprints@mathematik.uni-stuttgart.de
WWW: <http://preprints.ians.uni-stuttgart.de>

ISSN **1611-4176**

© Alle Rechte vorbehalten. Nachdruck nur mit Genehmigung des Autors.
IANS-Logo: Andreas Klimke. \LaTeX -Style: Winfried Geis, Thomas Merkle.

REDUZIERTE-BASIS-METHODEN

VORLESUNG VOM SS 2011

BERNARD HAASDONK

Institut für Angewandte Analysis und Numerische Simulation
Fakultät Mathematik und Physik
Universität Stuttgart

Stand: 26. Juli 2011

Vorwort

Dieses Skript beruht auf meiner Vorlesung *Reduzierte-Basis-Methoden* vom Sommersemester 2011, 2010 und 2009 an der Universität Stuttgart. Zentraler Inhalt dieser Vorlesung sind Reduzierte-Basis-Methoden zur Modellreduktion von parametrisierten elliptischen Differentialgleichungen.

Die Vorlesung wird durch 14-tägig ausgegebene Übungsblätter ergänzt. Die gesamte Veranstaltung ist eine gute Vorbereitung auf Abschluss-Arbeiten oder studentische Hilfskraft-Tätigkeiten im Rahmen aktueller Forschungsprojekte. Weitere Veranstaltungen (Vorlesung Modellreduktion, Seminar Angewandte Mathematik) eignen sich zur Ergänzung des Stoffes zu einem Prüfungsblock in Abschlussprüfungen.

Ich bedanke mich bei Herrn Sebastian Langhof für die Mitarbeit bei der Erstellung der initialen Form dieses Skriptes.

Bernard Haasdonk, Stuttgart 25. April 2011

Inhaltsverzeichnis

1	Einleitung	1
1.1	Parameterabhängige Probleme	1
1.2	Modellreduktion	4
1.3	Organisation der Vorlesung	5
2	Grundlagen	9
2.1	Lineare Funktionalanalysis in Hilberträumen	9
2.2	Bilinearformen	13
2.3	Parameterabhängigkeit	17
3	RB-Methoden für lineare koerzive Probleme	23
3.1	RB-Verfahren	23
3.2	Offline/Online-Zerlegung des RB-Modells	28
3.3	Fehleranalyse	29
3.4	Offline/Online-Zerlegung der Fehlerschätzer	37
4	A-priori Approximationstheorie	43
4.1	Approximierbarkeit durch lineare Unterräume	43
4.2	Fehlerfreie Approximationen	44
4.3	Asymptotische Approximationsaussagen	45
5	Basiskonstruktion	49
5.1	Motivation für Snapshot-basierte Verfahren	49
5.2	Gram-Schmidt	50
5.3	Proper Orthogonal Decomposition (POD)	52
5.4	Greedy Search	58
6	Allgemeine Parameterabhängigkeit	63
6.1	Empirische Interpolation (EI)	63
6.2	Eigenschaften der EI	67

6.3	Anwendung für lineare RB-Methoden	70
7	Allgemeine lineare Probleme	73
7.1	RB-Verfahren für allgemeine koerzive Probleme	73
7.2	Inf-sup-Stabilität	77
7.3	RB-Verfahren für inf-sup-stabile Probleme	80
8	Weiterführende Aspekte	85
8.1	Successive Constraints Method	85
8.2	Nichtlineare Probleme	87
	Literaturverzeichnis	90

Abbildungsverzeichnis

1.1	Beispiel Wärmeleitung mit Quelle Ω_q und Messbereich Ω_s . . .	2
1.2	Parametrisierte niedrigdimensionale Lösungsmenge	3
1.3	Beispiel einer interaktiven Oberfläche zur Parametervariation .	6
2.1	Orthogonale Projektion	11
2.2	Illustration zum Riesz'schen Darstellungssatz.	12
2.3	Komposit-Material	21
3.1	Laufzeitverhalten mit wachsender Anzahl von Simulationen . .	29
3.2	Fehlerabfall mit wachsender reduzierter Dimension.	32
5.1	POD: Ellipsoide aus Kovarianzoperator	53
6.1	EI von 3 Polynomen.	66
6.2	Empirische Interpolation für Polynome	67
8.1	Mengen der Successive Constraint Method	86

Kapitel 1

Einleitung

1.1 Parameterabhängige Probleme

(A) Parameterabhängige PDE

Sei $\Omega \subset \mathbb{R}^d$ polygonales Gebiet. Zu einem Parametervektor $\mu \in \mathcal{P} \subset \mathbb{R}^p$ aus einer Menge \mathcal{P} von „erlaubten“ Parametern ist eine Funktion, z.B. „Temperatur“

$$u(\mu) : \Omega \longrightarrow \mathbb{R}$$

gesucht, so dass

$$\begin{aligned} -\nabla \cdot (a(\mu)\nabla u) &= q(\mu) && \text{in } \Omega \\ u(\mu) &= 0 && \text{auf } \partial\Omega \end{aligned}$$

mit $a(\mu) : \Omega \longrightarrow \mathbb{R}$ „Wärmeleitkoeffizient“ und $q(\mu)$ eine „Wärmequelle“, z.B.:

$$q(x; \mu) = \begin{cases} 1 & \text{für } x \in \Omega_q \\ 0 & \text{sonst.} \end{cases}$$

Weiter kann eine Ausgabe erwünscht sein, z.B.:

$$s(\mu) = \frac{1}{|\Omega_S|} \int_{\Omega_S} u(x; \mu) \, dx$$

die „mittlere Temperatur auf Ω_S “, vgl. 1.1.

(B) Parameterabhängiges stationäres System

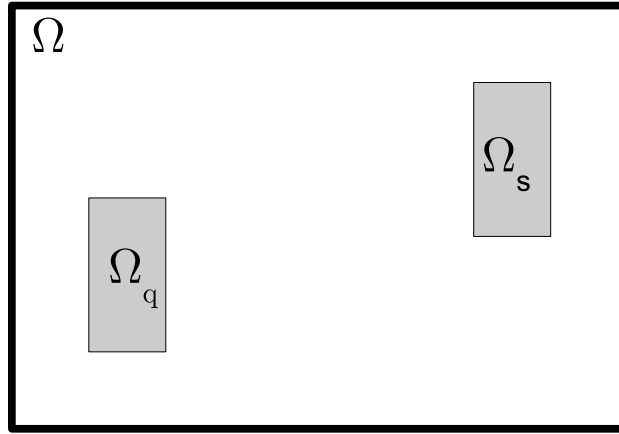


Abbildung 1.1: Beispiel Wärmeleitung mit Quelle Ω_q und Messbereich Ω_s

Zu einem Parametervektor $\mu \in \mathcal{P} \subset \mathbb{R}^P$ ist ein Zustandsvektor $u(\mu) \in \mathbb{R}^n$ und eine Ausgabe $s(\mu) \in \mathbb{R}^k$ gesucht, s.d.

$$\begin{aligned} 0 &= A(\mu)u(\mu) + B(\mu)w(\mu) \\ s(\mu) &= C(\mu)u(\mu) \end{aligned}$$

mit parameterabhängigen Matrizen $A(\mu) \in \mathbb{R}^{n \times n}$, $B(\mu) \in \mathbb{R}^{n \times m}$, $C(\mu) \in \mathbb{R}^{k \times n}$ und Eingangsvektor $w(\mu) \in \mathbb{R}^m$.

Schwache Formulierung in Hilberträumen

Sei X reeller Hilbertraum. Zu $\mu \in \mathcal{P}$ ist gesucht ein $u(\mu) \in X$ und Ausgabe $s(\mu) \in \mathbb{R}$ so, dass

$$b(u(\mu), v; \mu) = f(v; \mu), \quad s(\mu) = l(u(\mu); \mu) \quad \forall v \in X$$

mit Bilinearform $b(\cdot, \cdot; \mu)$ und linearem $f(\cdot; \mu), l(\cdot; \mu) : X \rightarrow \mathbb{R}$.

Beide oben angegebenen Beispiele lassen sich so formulieren:

$$\begin{aligned} \text{(A)} \quad X &:= H_0^1(\Omega) = \left\{ f \in L^2(\Omega) \mid \frac{\partial}{\partial x_i} f \in L^2(\Omega), f|_{\partial\Omega} = 0 \right\} \\ &\underbrace{\int_{\Omega} a(\mu) \nabla u(\mu) \cdot \nabla v}_{b(u(\mu), v; \mu)} = \underbrace{\int_{\Omega} q(\mu) v}_{f(v; \mu)} \quad \forall v \in X \\ s(\mu) &= \frac{1}{|\Omega_s|} \underbrace{\int_{\Omega_s(\mu)} u(x; \mu) dx}_{l(u(\mu); \mu)} \end{aligned}$$

(B) $X := \mathbb{R}^n$ ($k = 1$, „single input“)

$$\underbrace{v^T A(\mu) u(\mu)}_{b(u(\mu); v; \mu)} = \underbrace{-v^T B(\mu) w(\mu)}_{f(v; \mu)} \quad \forall v \in X$$

$$s(\mu) := \underbrace{C(\mu) u(\mu)}_{l(u(\mu); \mu)}$$

In der Vorlesung werden weitere Verallgemeinerungen zu $b : X_1 \times X_2 \rightarrow \mathbb{R}$ mit $X_1 \neq X_2$ und Nichtlinearitäten behandelt.

Parameterabhängige Lösungsmenge

$\mathcal{M} := \{u(\mu) \mid \mu \in \mathcal{P}\} \subset X$ für $\mathcal{P} \subset \mathbb{R}^p$ ist die durch μ parametrisierte Lösungsmenge.

X ist im allgemeinen unendlich dimensional (PDE) oder hochdimensional (FEM, FV-Raum), \mathcal{M} ist jedoch höchstens p -dimensional.

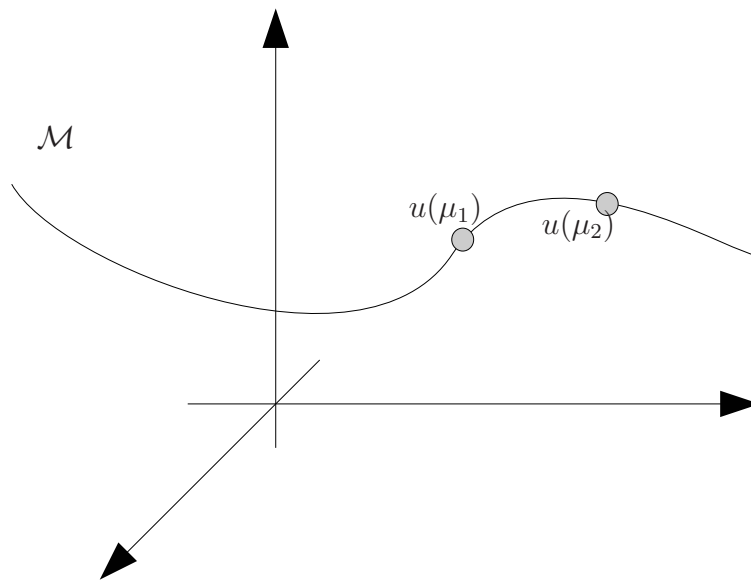


Abbildung 1.2: Parametrisierte niedrigdimensionale Lösungsmenge

\Rightarrow Motivation für die Suche nach einem „niedrigdimensionalen“ Teilraum $X_N \subset X$ zur Approximation von \mathcal{M} und einer Approximation $u_N(\mu) \approx u(\mu)$, $u_N \in X_N$.

Insbesondere Idee bei Reduzierte-Basis-Methoden:

X_N durch Beispiellösungen erzeugt, sogenannte „Snapshots“.

$X_N \subseteq \text{span}\{u(\mu_1), \dots, u(\mu_n)\}$ für geeignete Parameter $\mu_1, \dots, \mu_n \in \mathcal{P}$.

Beispiel

Gesucht ist $u(\cdot; \mu) \in C^2([0, 1])$ mit

$$\begin{aligned}(1 + \mu)u'' &= 1 \text{ auf } (0, 1) \\ u(0) &= u(1) = 1\end{aligned}$$

für Parameter $\mu \in \mathcal{P} := [0, 1] \subset \mathbb{R}$.

Spezielle Lösungen „Snapshots“:

$$\begin{aligned}\mu = 0 &\Rightarrow u_0 = u(\cdot; \mu = 0) = \frac{1}{2}x^2 - \frac{1}{2}x + 1 \\ \mu = 1 &\Rightarrow u_1 = u(\cdot; \mu = 1) = \frac{1}{4}x^2 - \frac{1}{4}x + 1 \\ X_N &:= \text{span}\{u_0, u_1\}\end{aligned}$$

Reduzierte Lösung $u_N(\mu) \in X_N$ gegeben durch:

$$\begin{aligned}u_N(\mu) &= \alpha_0(\mu)u_0 + \alpha_1(\mu)u_1 \\ \alpha_0(\mu) &= \frac{2}{\mu + 1} - 1, \quad \alpha_1(\mu) = 2 - \frac{2}{\mu + 1}\end{aligned}$$

Diese erfüllt folgende Fehleraussage und ist somit exakt:

$$\|u_N(\mu) - u(\mu)\|_\infty = \sup_{x \in [0, 1]} |u_N(x, \mu) - u(x; \mu)| = 0.$$

M ist enthalten im 2-dimensionalen Raum X_N . Genauer: $\alpha_0 + \alpha_1 = 1, 0 \leq \alpha_0, \alpha_1 \leq 1$, also ist \mathcal{M} Menge der Konvexkombinationen von u_0, u_1 .

1.2 Modellreduktion

Begrifflichkeiten

- Eine PDE ist ein *analytisches Modell*, welches die *analytische Lösung* $u(\mu) \in X$ in einem (typischerweise ∞ -dimensionalen) Funktionenraum charakterisiert.
- Ein *detailliertes Modell* ist ein Berechnungsverfahren zur Bestimmung einer Näherung $u(\mu) \in X$ in einem hochdimensionalen Funktionenraum

mit sehr allgemeinen Approximationseigenschaften (Finite Elemente, Finite Differenzen, Finite Volumen, $\dim X$ z.B. $10^3 - 10^8$). Somit kann $u(\mu)$ also sowohl ein analytisches als auch detailliertes Modell bezeichnen.

- Ein *reduziertes Modell* ist ein Berechnungsverfahren zur Bestimmung einer Näherung $u_N(\mu) \in X_N$ in einem sehr problemangepassten und daher niedrigdimensionalen Raum ($\dim X$ z.B. $1, \dots, 100$).
- *Modellreduktion* beschäftigt sich mit Methoden der Erzeugung von reduzierten Modellen aus detaillierten (oder auch analytischen) Modellen und Untersuchungen ihrer Eigenschaften.

Modellreduktion ist ein modernes und zukunftsträchtiges Gebiet der Angewandten Mathematik und Ingenieurwissenschaften. (Workshop MoRePaS 09, SimTech-Cluster: PN6-Seminar SS 2011, Di. 11:30-13:00, SR 7.122)

Anwendungen für parametrisierte reduzierte Modelle

Vielfach-Anfragen („multi-query“) unter Parametervariation:

- Parameterstudien
- Design
- Parameteridentifikation/Inverse Probleme
- Optimierung
- Statistische Analyse

Schnelle Simulationsantwort:

- Echtzeit-Steuerung technischer Geräte
- Webformulare
- Interaktive Benutzeroberflächen
- Smartphone-Apps

Demonstration: Online-GUI, Handy-App

1.3 Organisation der Vorlesung

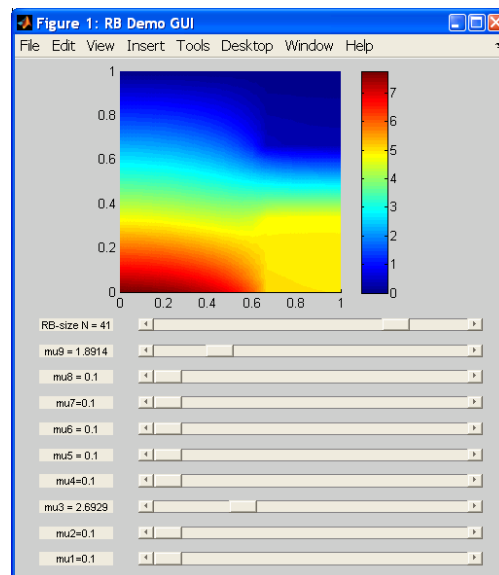


Abbildung 1.3: Beispiel einer interaktiven Oberfläche zur Parametervariation

Zentrale Fragen

- Reduzierte Basis: Wie kann ein möglichst kompakter Teilraum konstruiert werden?
- Reduziertes Modell: Wie kann eine reduzierte Lösung $u_N(\mu)$ berechnet werden?
- Effizienz: Wie kann $u_N(\mu)$ schnell berechnet werden?
- Stabilität: Wie kann die Stabilität des reduzierten Modells für wachsendes N garantiert werden?
- Fehlerschätzer: Kann der Fehler des reduzierten zum vollen Modell beschränkt werden?
- Effektivität von Fehlerschätzer: Kann garantiert werden, dass Fehlerschätzer den tatsächlichen Fehler nicht beliebig überschätzen?

Vorläufige Gliederung

- 1) Einleitung
- 2) Grundlagen
- 3) RB-Verfahren für lineare koerzive Probleme

- 4) A-priori Theorie
- 5) Basis Generierung
- 6) Allgemeine Parameterabhängigkeit
- 7) Allgemeine lineare Probleme
- 8) Weiterführende Aspekte

Die Vorlesung „Modellreduktion“ WS 09/10 oder das Seminar „Angewandte Mathematik“ können als Ergänzungen dieser Veranstaltung gesehen werden.

Referenzen

Literatur

- Allgemeine Darstellung von RB-Methoden für parametrisierte elliptische Probleme: E-Book [PR07]
A.T. Patera, G. Rozza: Reduced Basis Approximation and A posteriori Error Estimation for Parametrized Partial Differential Equations, MIT 2006-2007.
- Modellreduktion für Dynamische Systeme: [Ant05]
A.C. Antoulas: Approximation of Large-Scale Dynamical Systems, SIAM 2005
- Funktionalanalysis: [Alt92]
H.W. Alt: Lineare Funktionalanalysis, Springer 1992
- Finite Elemente: [Bra03]
D. Braess: Finite Elemente, Springer 2003

Webseiten

- <http://www.morepas.org>
- <http://augustine.mit.edu/methodology.htm>

Weitere spezielle Literaturhinweise folgen an geeigneten Stellen.

Kapitel 2

Grundlagen

2.1 Lineare Funktionalanalysis in Hilberträumen

Definition 2.1 (Hilbertraum):

Sei X ein reeller Vektorraum mit $(\cdot, \cdot) : X \times X \rightarrow \mathbb{R}$ ein Skalarprodukt und induzierter Norm $\|x\| := \sqrt{(x, x)}$. Falls X vollständig bzgl. $\|\cdot\|$, ist es ein (reeller) Hilbertraum (HR).

Beispiel (siehe Übung):

- i) $X := \mathbb{R}^d$ mit $(x, x') := \sum_{i=1}^d x_i x'_i$ ist HR.
- ii) $X := C^0([0, 1])$ mit $(f, g) := \int_0^1 f(x)g(x)dx$ ist kein HR.
- iii) $X := H_0^1(\Omega)$ mit $(f, g) := \int_{\Omega} \nabla f(x) \cdot \nabla g(x)dx$ ist HR.
- iv) $X := H^1(\Omega)$ mit $(f, g) := \int_{\Omega} f(x)g(x)dx + \int_{\Omega} \nabla f(x) \cdot \nabla g(x)dx$ ist HR.
- v) Seien $(X_1, (\cdot, \cdot)_1)$ und $(X_2, (\cdot, \cdot)_2)$ HR. Dann ist $X := X_1 \times X_2$ mit $(f, g) := (f_1, g_1)_1 + (f_2, g_2)_2$ für $f = (f_1, f_2), g = (g_1, g_2) \in X_1 \times X_2$ der Produktraum von X_1 und X_2 .

Definition 2.2 (Lineare Operatoren):

Seien X, Y HR (reelle Hilberträume)

- i) Eine lineare Abbildung $A : X \rightarrow Y$ ist ein stetiger linearer Operator, falls

$$\|A\| := \sup_{x \in X \setminus \{0\}} \frac{\|Ax\|}{\|x\|} < \infty.$$

ii) $L(X, Y)$ ist der Raum der stetigen linearen Operatoren von X nach Y . $L(X) := L(X, X)$.

iii) $X' := L(X, \mathbb{R})$ ist der Dualraum von X , die Elemente sind stetige lineare Funktionale.

iv) Für $A \in L(X, Y)$ bezeichnen wir

$$\begin{aligned} \text{Bild}(A) &:= A(X) \subset Y \text{ die } \underline{\text{Bildmenge}} \\ \text{Kern}(A) &:= A^{-1}(0) \subset X \text{ der } \underline{\text{Kern}}. \end{aligned}$$

v) Wir bezeichnen mit $B_1(0) \subset X$ die offene Einheitskugel. Dann ist $K(X, Y) := \{A \in L(X, Y) \mid A(B_1(0)) \text{ kompakt}\}$ der Raum der kompakten Operatoren,

vi) $P \in L(X)$ ist eine Projektion $:\Leftrightarrow P^2 := P \circ P = P$.

vii) $A \in L(X, Y)$ ist Isometrie, falls $\|x\| = \|Ax\| \quad \forall x \in X$.

Satz 2.3 (Orthogonale Projektion):

Sei X HR, $Y \subset X$ abgeschlossener Teilraum

i) Es existiert ein eindeutiges $P \in L(X, Y)$ mit

$$(x - Px, y) = 0 \quad \forall x \in X, y \in Y.$$

ii) P erfüllt

$$\|x - Px\| = \inf_{y \in Y} \|x - y\| \quad \forall x \in X.$$

iii) Falls $\dim Y = n < \infty$ und $\{\varphi_i\}_{i=1}^n$ ist Orthonormalbasis von Y , gilt

$$Px = \sum_{i=1}^n (x, \varphi_i) \varphi_i.$$

Beweis: Übung □

Satz 2.4 (Riesz'scher Darstellungssatz):

Sei X ein HR, so ist $J : X \rightarrow X'$ definiert durch $J(v)(w) := (v, w) \quad \forall v, w \in X$ eine stetige, lineare, bijektive Isometrie. Insbesondere existiert zu $l \in X'$ ein eindeutiger Riesz Repräsentant $v_l := J^{-1}(l) \in X$ mit $l(\cdot) = (v_l, \cdot)$.

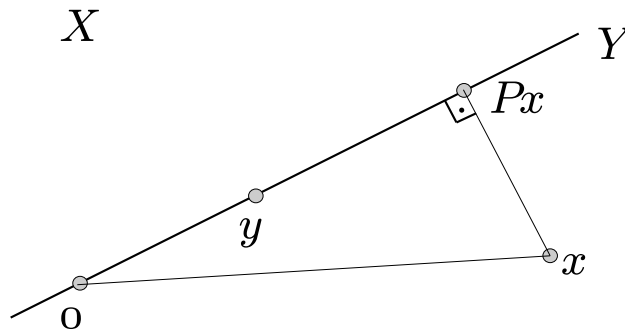


Abbildung 2.1: Orthogonale Projektion

Beweis: J ist linear:

$$\begin{aligned} J(\lambda_1 v_1 + \lambda_2 v_2)(\cdot) &= (\lambda_1 v_1 + \lambda_2 v_2, \cdot) \\ &= \lambda_1 (v_1, \cdot) + \lambda_2 (v_2, \cdot) \\ &= \lambda_1 J(v_1)(\cdot) + \lambda_2 J(v_2)(\cdot) \end{aligned}$$

Sei $v \in X \setminus \{0\}$

$$\begin{aligned} \|J(v)\| &= \sup_{w \in X \setminus \{0\}} \frac{|J(v)(w)|}{\|w\|} = \sup_w \frac{|(v, w)|}{\|w\|} \\ &\stackrel{CSU}{=} \frac{\|v\| \|v\|}{\|v\|} = \|v\| \\ &\implies \|J\| = 1 \end{aligned}$$

also ist J stetig, Isometrie. $\text{Kern}(J) = \{0\}$, also J injektiv.

Zeige J surjektiv: Sei $l \in X', l \neq 0$. $\text{Kern}(l)$ ist abgeschlossener Teilraum, also existiert $P : X \rightarrow \text{Kern}(l)$ orthogonale Projektion nach Satz 2.3. Sei $v_0 \in X$ mit $l(v_0) = 1$. Setze $v_1 := v_0 - Pv_0 \Rightarrow l(v_1) = l(v_0) = 1, v_1 \perp \text{Kern}(l)$. Für $v \in X$ gilt

$$v = \underbrace{v - l(v) \cdot v_1}_{\in \text{Kern}(l)} + l(v) \cdot v_1$$

und $v - l(v)v_1 \in \text{Kern}(l)$ wegen $l(v - l(v)v_1) = l(v) - l(v)l(v_1) = 0$. Also ist

$$\begin{aligned} \left(\frac{v_1}{\|v_1\|^2}, v \right) &= \underbrace{\left(\frac{v_1}{\|v_1\|^2}, v - l(v) \cdot v_1 \right)}_{=0, \text{ da } v_1 \perp \text{Kern}(l)} + \left(\frac{v_1}{\|v_1\|^2}, l(v) \cdot v_1 \right) \\ &= l(v) \cdot \frac{\overrightarrow{(v_1, v_1)}}{\|v_1\|^2} = l(v). \end{aligned}$$

Also $l \in \text{Bild}(J)$, J bijektiv.

Damit erfüllt $v_l := \frac{v_1}{\|v_1\|^2}$ das Gewünschte. \square

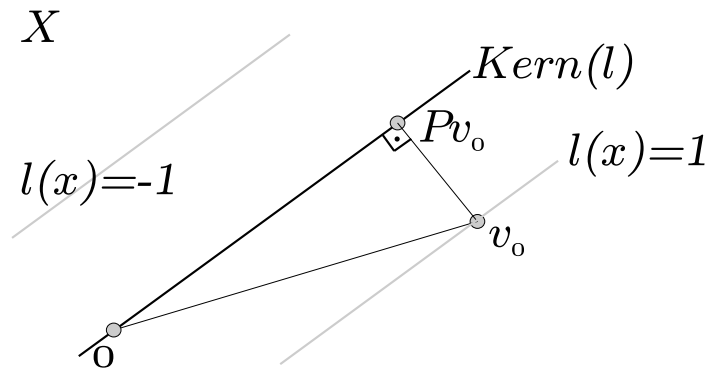


Abbildung 2.2: Illustration zum Riesz'schen Darstellungssatz.

Satz 2.5 (Adjungierter Operator):

Seien X, Y HR. Zu $A \in L(X, Y)$ existiert ein eindeutiger Operator $A^* \in L(Y, X)$ mit

$$(Ax, y) = (x, A^*y) \quad \forall x \in X, y \in Y$$

der sogenannte adjungierte Operator. Falls $X = Y$ und $A = A^*$, so nennt man A selbstadjungiert.

Beweis: Für $y \in Y$ ist $(A \cdot, y) \in X'$, also existiert ein eindeutiger Riesz Repräsentant $A^*y \in X$ mit $(A \cdot, y) = (\cdot, A^*y)$, $A^* : Y \rightarrow X$ also eindeutig definiert.

Linearität:

$$\begin{aligned} (x, A^*(\lambda_1 y_1 + \lambda_2 y_2)) &\stackrel{Def}{=} (Ax, \lambda_1 y_1 + \lambda_2 y_2) = \lambda_1 (Ax, y_1) + \lambda_2 (Ax, y_2) \\ &= \lambda_1 (x, A^*y_1) + \lambda_2 (x, A^*y_2) = (x, \lambda_1 A^*y_1 + \lambda_2 A^*y_2) \end{aligned}$$

Stetigkeit:

$$\begin{aligned} \|A^*y\| &= \frac{\|A^*y\|^2}{\|A^*y\|} \stackrel{CSU}{=} \sup_{x \in X \setminus \{0\}} \frac{(x, A^*y)}{\|x\|} = \sup_{x \in X \setminus \{0\}} \frac{(Ax, y)}{\|x\|} \\ &\stackrel{CSU}{\leq} \sup_x \frac{\|A\| \|x\| \|y\|}{\|x\|} = \|A\| \|y\| \\ &\Rightarrow \|A^*\| \leq \|A\| \end{aligned}$$

also $\|A^*\|$ stetig, da $\|A\|$ stetig. \square

Satz 2.6 (Spektralsatz für kompakte selbstadjungierte Operatoren):
Sei X HR, $A \in K(X, X)$ selbstadjungiert, dann existiert ein endliches oder abzählbar unendliches Orthonormalsystem von Eigenvektoren $\{\varphi_i\}_{i \in I}$, $I \subset \mathbb{N}$ zu Eigenwerten $\{\lambda_i\}_{i \in I} \subset \mathbb{R} \setminus \{0\}$ mit

$$Ax = \sum_{i \in I} \lambda_i(x, \varphi_i) \varphi_i \quad \forall x \in X.$$

Falls I unendlich, so $\lim_{i \rightarrow \infty} \lambda_i = 0$.

Beweis: \Rightarrow [Alt92], Satz 10.12, S. 300. □

2.2 Bilinearformen

Definition 2.7 (Bilinearformen):

Seien X_1, X_2 HR, $b : X_1 \times X_2 \rightarrow \mathbb{R}$ Bilinearform.

i) Falls

$$\gamma := \sup_{u \in X_1 \setminus \{0\}} \sup_{v \in X_2 \setminus \{0\}} \frac{b(u, v)}{\|u\| \|v\|} < \infty,$$

so ist b stetig mit Stetigkeitskonstante γ .

ii) Falls $X = X_1 = X_2$, definieren wir

$$\begin{aligned} b_s(u, v) &:= \frac{1}{2}(b(u, v) + b(v, u)) \\ b_a(u, v) &:= \frac{1}{2}(b(u, v) - b(v, u)) \quad \forall u, v \in X \end{aligned}$$

den symmetrischen bzw. antisymmetrischen Anteil von $b = b_s + b_a$.

iii) Falls $X = X_1 = X_2$, b stetig und

$$\alpha := \inf_{u \in X \setminus \{0\}} \frac{b(u, u)}{\|u\|^2} > 0$$

So heißt b koerziv mit Koerzivitätskonstante α .

Bemerkung:

i) $\alpha \in \mathbb{R}$ ist wohldefiniert, denn mit Stetigkeit nach unten beschränkt:

$$\frac{b(u, u)}{\|u\|^2} \geq -\gamma \frac{\|u\| \|u\|}{\|u\|^2} = -\gamma$$

ii) b ist koerziv bzgl. $\alpha \Leftrightarrow b_s$ koerziv bzgl. α .

$$\text{Denn } b(u, u) = b_s(u, u) + \underbrace{b_a(u, u)}_{=0} = b_s(u, u)$$

$$\text{also } \alpha \|u\|^2 \leq b(u, u) \Leftrightarrow \alpha \|u\|^2 \leq b_s(u, u).$$

iii) Mit der Stetigkeit von b sind auch b_s und b_a stetig.

iv) Es gilt b koerziv $\Leftrightarrow \lim_{u \rightarrow \infty} b(u, u) = \infty$.

Satz 2.8 (Operatoren und Bilinearformen):

Seien X_1, X_2 HR.

i) Zu $B \in L(X_1, X_2)$ existiert eine eindeutig definierte stetige Bilinearform $b : X_1 \times X_2 \rightarrow \mathbb{R}$ mit

$$b(u, v) = (Bu, v) \quad \forall u \in X_1, v \in X_2. \quad (2.1)$$

ii) Zu $b : X_1 \times X_2 \rightarrow \mathbb{R}$ stetige Bilinearform existiert eindeutiges $B \in L(X_1, X_2)$ welches (2.1) erfüllt.

Beweis:

i) b definiert durch (2.1) ist bilinear wegen Bilinearität von (\cdot, \cdot) und Linearität von B .

Stetigkeit:

$$\begin{aligned} b(u, v) &= (Bu, v) \stackrel{CSU}{\leq} \|B\| \|u\| \|v\| \\ &\Rightarrow \gamma \leq \|B\| < \infty \end{aligned}$$

ii) Sei $u \in X_1$ fest, dann ist $b(u, \cdot) : X_2 \rightarrow \mathbb{R}$ linear und stetig:

$$\sup_{v \in X_2} \frac{b(u, v)}{\|v\|} \leq \sup_v \gamma \frac{\|u\| \|v\|}{\|v\|} = \gamma \|u\| < \infty$$

$\Rightarrow b(u, \cdot) \in X_2'$ und es existiert nach Satz 2.3 ein eindeutiger Riesz-Repräsentant $v_u \in X_2$ mit $b(u, \cdot) = (v_u, \cdot)$. Definiere $B : X_1 \rightarrow X_2$ durch $Bu := v_u \in X_2$. Hiermit (2.1) und Eindeutigkeit klar.

Linearität: standard

Stetigkeit:

$$\begin{aligned} \|Bu\|^2 &\stackrel{Def}{=} (Bu, Bu) \stackrel{Def}{=} (v_u, Bu) \\ &= b(u, Bu) \leq \gamma \|u\| \|Bu\| \\ &\Rightarrow \|Bu\| \leq \gamma \|u\| \\ &\Rightarrow \sup_u \frac{\|Bu\|}{\|u\|} \leq \gamma. \end{aligned}$$

□

Satz 2.9 (Eigenwertproblem für α):

Sei X HR, $b : X \times X$ stetige, koerzive Bilinearform, der symmetrische Anteil b_s repräsentiert durch $B_s \in K(X, X)$

$$b_s(u, v) = (B_s u, v) \quad \forall u, v \in X$$

Dann ist

$$\alpha := \inf_{u \in X} \frac{b_s(u, u)}{\|u\|^2}$$

der kleinste Eigenwert des verallgemeinerten Eigenwertproblems:

$$\text{gesucht } (u, \lambda) \in X \times \mathbb{R} \quad \text{s.d.} \quad b_s(u, v) = \lambda \cdot (u, v) \quad \forall v \in X \quad (2.2)$$

Beweis: Sei $(u, \lambda) \in X \times \mathbb{R}$ Eigenvektor/Eigenwert. Da B_s kompakt und selbstadjungiert existiert nach Satz 2.6 folgende Darstellung

$$B_s u = \sum_{i \in I} \lambda_i(u, \varphi_i) \varphi_i,$$

$I \subset \mathbb{N}$, $\{\varphi_i\}$ orthonormale Eigenvektoren, $\lambda_i \neq 0$.

Annahme: I unendlich. Mit Spektralsatz gilt dann $\lambda_i \rightarrow 0$, also

$$\frac{b_s(\varphi_i, \varphi_i)}{\|\varphi_i\|^2} = \frac{\lambda_i \|\varphi_i\|^2}{\|\varphi_i\|^2} \rightarrow 0.$$

⚡ Widerspruch zu Koerzivität.

Also I endlich, o.B.d.A. $I = \{1, \dots, n\}$. Sei $u \in X$, $u = \varphi_0 + \sum_{i=1}^n u_i \varphi_i$ für geeignete $u_i \in \mathbb{R}$ und $\varphi_0 \perp \text{Bild}(B)$. Dann ist $\varphi_0 = 0$ (sonst $B\varphi_0 = 0 = \lambda_0 \varphi_0$ also EV zu $\lambda = 0$, Widerspruch zu Koerzivität), o.B.d.A. $\|u\|^2 = 1$ also $\sum u_i^2 = 1$ (denn $\|u\|^2 = \sum u_i^2$). Dann gilt

$$\frac{b_s(u, u)}{\|u\|^2} = (B_s u, u) = \sum_{i=1}^n \lambda_i u_i^2$$

also Konvexkombination der λ_i mit Koeffizienten u_i^2 (denn $0 \leq u_i^2 \leq 1$, $\sum_i u_i^2 = 1$). Damit ist

$$\alpha = \inf_u \sum_{i=1}^n \lambda_i u_i^2 \geq \min \lambda_i.$$

Sei $u = \varphi_i$ mit $\lambda_i = \min_{j=1, \dots, n} \lambda_j$

$$\begin{aligned} \frac{b_s(u, u)}{\|u\|^2} &= \lambda_i = \min_{j=1, \dots, n} \lambda_j \\ \Rightarrow \alpha &= \min \lambda_i \end{aligned}$$

□

Satz 2.10 (Koerzivität in \mathbb{R}^n):

Sei $X := \mathbb{R}^n$, $b(u, v) := (Bu, v)$ für $B \in \mathbb{R}^{n \times n}$ und $\alpha := \inf_{u \in X \setminus \{0\}} \frac{b(u, u)}{\|u\|^2}$. Seien $\{\lambda_i\}_{i=1}^n$ die Eigenwerte von $B_s := \frac{1}{2}(B + B^T)$. Dann gilt

i) $\alpha = \min_{i=1, \dots, n} \lambda_i$

ii) b ist koerziv $\Leftrightarrow B_s$ ist positiv definit.

Beweis: Übung. □

Satz 2.11 (Lax-Milgram):

. Sei X HR, $b : X \times X \rightarrow \mathbb{R}$ koerzive, stetige Bilinearform mit Koerzivitätskonstante α . Dann existiert ein eindeutiger Operator $B \in L(X)$ mit

$$b(u, v) = (Bu, v) \quad \forall u, v \in X,$$

B ist bijektiv, $B^{-1} \in L(X)$ und

$$\|B^{-1}\| \leq \frac{1}{\alpha}.$$

Beweis: Existenz und Eindeutigkeit von B klar nach Satz 2.8. Mit Koerzivität erhalten wir

$$\begin{aligned} \alpha \|u\|^2 \leq b(u, u) &= (Bu, u) \stackrel{CSU}{\leq} \|Bu\| \|u\| \\ \|Bu\| &\geq \alpha \|u\| \quad \forall u \in X \end{aligned} \tag{2.3}$$

Injektivität:

$$Bu = 0 \stackrel{(2.3)}{\Rightarrow} \|u\| = 0 \Rightarrow u = 0$$

Zeige zunächst Abgeschlossenheit von $\text{Bild}(B)$: Sei $(Bu)_{i \in \mathbb{N}}$ Cauchyfolge im $\text{Bild}(B)$. Mit (2.3) folgt

$$\|Bu_m - Bu_n\| \geq \alpha \|u_m - u_n\|$$

$\Rightarrow (u_i)_{i \in \mathbb{N}}$ ist Cauchyfolge, es existiert $u = \lim_{i \rightarrow \infty} u_i \in X$ wegen Vollständigkeit. Mit Stetigkeit von B folgt

$$\lim_{i \rightarrow \infty} Bu_i = Bu \in \text{Bild}(B)$$

$\Rightarrow \text{Bild}(B)$ abgeschlossen. Damit existiert nach Satz 2.3 orthogonale Projektion P auf $\text{Bild}(B)$.

Surjektivität: Sei $v \in X, \bar{v} := v - Pv$. Dann ist $(Bu, \bar{v}) = 0 \quad \forall u \in X$ und

$$\begin{aligned} \alpha \|\bar{v}\|^2 &\leq b(\bar{v}, \bar{v}) = (B\bar{v}, \bar{v}) = 0 \\ &\Rightarrow \bar{v} = 0 \\ &\Rightarrow v = Pv \in \text{Bild}(B). \end{aligned}$$

d.h. B surjektiv, also bijektiv. Somit existiert B^{-1} linear. Stetigkeit folgt mit $v := Bu$ in (2.3):

$$\begin{aligned} \|v\| &\geq \alpha \|u\| = \alpha \|B^{-1}v\| \\ \Rightarrow \|B^{-1}v\| &\leq \frac{1}{\alpha} \|v\| \\ \Rightarrow \|B^{-1}\| &= \sup_{v \neq 0} \frac{\|B^{-1}v\|}{\|v\|} \leq \frac{1}{\alpha}. \end{aligned}$$

□

2.3 Parameterabhängigkeit

Definition 2.12 (Parametrische Formen):

Sei $\mathcal{P} \subset \mathbb{R}^p$ eine beschränkte Parametermenge. Dann nennen wir

i) $l : X \times \mathcal{P} \rightarrow \mathbb{R}$ eine parametrische stetige Linearform falls $\forall \mu \in \mathcal{P}$

$$l(\cdot; \mu) \in X'.$$

ii) Wir nennen $b : X_1 \times X_2 \times \mathcal{P} \rightarrow \mathbb{R}$ eine parametrische stetige (symmetrische / koerzive) Bilinearform, falls $\forall \mu \in \mathcal{P}$

$$b(\cdot, \cdot; \mu) : X_1 \times X_2 \rightarrow \mathbb{R} \text{ bilinear (symmetrisch / koerziv)}$$

und bezeichnen die Stetigkeitskonstante mit $\gamma(\mu)$ und die Koerzivitätskonstante mit $\alpha(\mu)$.

Bemerkung:

Eine parametrische stetige Bi-/Linearform ist nicht unbedingt stetig bzgl. μ :
Beispiel:

$$X = \mathbb{R}, \mathcal{P} = [0, 1], l : X \times \mathcal{P} \rightarrow \mathbb{R}$$

$$l(x; \mu) := \begin{cases} x & , \text{ falls } \mu < \frac{1}{2} \\ \frac{1}{2}x & , \text{ sonst} \end{cases}.$$

Definition 2.13 (Parametrische Beschränktheit/Stetigkeit):

- i) Wir nennen eine parametrische stetige Linearform l bzw. Bilinearform b gleichmäßig beschränkt bzgl. μ falls existiert $\gamma_0 \in \mathbb{R}^+$ mit

$$\sup_{\mu \in \mathcal{P}} \|l(\cdot; \mu)\|_{X'} \leq \gamma_0, \quad \text{bzw.} \quad \sup_{\mu \in \mathcal{P}} \gamma(\mu) \leq \gamma_0.$$

- ii) Wir nennen b gleichmäßig koerziv bzgl. μ falls existiert $\alpha_0 > 0$ mit

$$\inf_{\mu \in \mathcal{P}} \alpha(\mu) \geq \alpha_0.$$

- iii) Wir nennen l bzw. b Lipschitz-stetig bzgl. μ falls existiert ein $L_l \in \mathbb{R}^+$ bzw. $L_b \in \mathbb{R}^+$ so dass für alle $\mu_1, \mu_2 \in \mathcal{P}$ gilt

$$\begin{aligned} |l(u; \mu_1) - l(u; \mu_2)| &< L_l \|u\| \|\mu_1 - \mu_2\| \quad \forall u \in X, \quad \text{bzw.} \\ |b(u, v; \mu_1) - b(u, v; \mu_2)| &< L_b \|u\| \|v\| \|\mu_1 - \mu_2\| \quad \forall u \in X_1, v \in X_2. \end{aligned}$$

Definition 2.14 (Sensitivitätsableitung):

Sei X HR, $\mathcal{P} \subset \mathbb{R}^p$ beschränkt, $\mu_0 \in \mathcal{U} \subset \mathcal{P}$ in Umgebung \mathcal{U} . $f : \mathcal{U} \rightarrow X$ ist (Fréchet)-differenzierbar in μ_0 , falls existiert ein $Df(\mu_0) \in L(\mathbb{R}^p, X)$ mit

$$\lim_{h \rightarrow 0} \frac{\|f(\mu_0 + h) - f(\mu_0) - Df(\mu_0)h\|}{\|h\|} = 0.$$

Falls f in jedem $\mu \in \mathcal{U}$ diff'bar, dann existieren insbesondere partielle Ableitungen

$$\frac{\partial}{\partial \mu_i} f(\cdot) := Df(\cdot)e_i : \mathcal{U} \rightarrow X$$

für $e_i \in \mathbb{R}^p$ Einheitsvektor $i = 1, \dots, p$.

Falls diese wiederum diff'bar in \mathcal{U} bezeichnet allgemein

$$\partial_\sigma f(\cdot) := \frac{\partial^{|\sigma|}}{\partial \mu_1^{\sigma_1} \dots \partial \mu_p^{\sigma_p}} f(\cdot) : \mathcal{U} \rightarrow X$$

die Sensitivitätsableitung der Ordnung $|\sigma| := \sum_{i=1}^p \sigma_i$ für Multiindex $\sigma = (\sigma_1, \dots, \sigma_p) \in \mathbb{N}_0^p$

Bemerkung:

Diese Ableitungen werden später insbesondere bei parameterabhängigen Lösungen $u(x; \mu)$ verwendet $u : \Omega \times \mathcal{P} \rightarrow \mathbb{R}$ mit $u(\cdot; \mu) \in X \quad \forall \mu$ kann auch als $u : \mathcal{P} \rightarrow X$ aufgefasst werden mit Sensitivitätsableiten $\partial_\sigma u : \mathcal{P} \rightarrow X$ d.h. $\partial_\sigma u(\cdot; \mu) \in X \quad \forall \mu \in \mathcal{P}$ und insbesondere $\partial_\sigma u : \Omega \times \mathcal{P} \rightarrow \mathbb{R}$ d.h. Funktionen $\partial_\sigma u(x; \mu)$

Definition 2.15 (Separierbare Parameterabhängigkeit):

Seien X, X_1, X_2 HR, \mathcal{P} beschränkte Parametermenge

- i) Eine Funktion $v : \mathcal{P} \rightarrow X$ nennen wir separierbar parametrisch, falls existieren Komponenten $v^q \in X$ und Koeffizientenfunktionen $\Theta_v^q : \mathcal{P} \rightarrow \mathbb{R}$ für $q = 1, \dots, Q_v$ mit

$$v(\mu) = \sum_{q=1}^{Q_v} \Theta_v^q(\mu) v^q$$

- ii) Eine parametrische stetige Linearform $l : X \times \mathcal{P} \rightarrow \mathbb{R}$ bzw. Bilinearform $b : X_1 \times X_2 \times \mathcal{P} \rightarrow \mathbb{R}$ ist separierbar parametrisch, falls existieren $l^q \in X'$ und $\Theta_l^q : \mathcal{P} \rightarrow \mathbb{R}$ für $q = 1, \dots, Q_l$ bzw. $b^q : X_1 \times X_2 \rightarrow \mathbb{R}$ stetig und $\Theta_b^q : \mathcal{P} \rightarrow \mathbb{R}$ für $q = 1, \dots, Q_b$ mit

$$l(v; \mu) = \sum_{q=1}^{Q_l} \Theta_l^q(\mu) l^q(v) \quad \forall v \in X \quad \text{bzw.}$$

$$b(u, v; \mu) = \sum_{q=1}^{Q_b} \Theta_b^q(\mu) b^q(u, v) \quad \forall (u, v) \in X_1 \times X_2.$$

Lemma 2.16 (Energienorm):

Sei X HR, $b : X \times X \times \mathcal{P} \rightarrow \mathbb{R}$ parametrische, koerzive, stetige Bilinearform. Dann ist für $\mu \in \mathcal{P}$

$$(((u, v)))_\mu := b_s(u, v; \mu)$$

ein Skalarprodukt auf X und

$$|||u|||_\mu := \sqrt{(((u, u)))_\mu}$$

die „Energienorm“. Diese ist äquivalent zur X -Norm

$$\sqrt{\alpha(\mu)} \|u\| \leq |||u|||_\mu \leq \sqrt{\gamma(\mu)} \|u\| \quad \forall u \in X.$$

Beweis: Skalarprodukt: klar, wegen Bilinearität, Stetigkeit und Koerzivität. Normäquivalenz folgt aus Stetigkeit und Koerzivität von b_s .

$$\alpha(\mu)\|u\|^2 \leq b_s(u, u; \mu) \leq \gamma(\mu)\|u\|^2$$

□

Beispiel 2.17 (Wärmeleitung in Komposit-Material):

Sei wie in Abb. 2.17:

- $\Omega = \bigcup_{i=1}^p \Omega_i$ Gebiet aus $p \in \mathbb{N}$ polygonalen Teilgebieten Ω_i
- $\mathcal{P} = [\mu_{min}, \mu_{max}]^p \subset \mathbb{R}^p$ Parameterbereich mit $0 < \mu_{min} < \mu_{max}$.
- $\mu = (\mu_1, \dots, \mu_p)^T \in \mathcal{P}$ Vektor aus Wärmeleitkoeffizienten der Teilgebiete, $a(x)$ ortsabhängiger, stückweise konstanter Wärmeleitkoeffizient.
- PDE: Wärmeleitung ohne innere Quelle, $-\nabla \cdot (a\nabla u) = 0$ in Ω
- Dirichlet 0 Randbedingung: $u = 0$ auf Γ_D („Kühlung auf Temperatur 0“)
- Neumann-Randbedingung ohne Fluss: $a\nabla u \cdot n = 0$ auf Γ_{N0} („isolierte Ränder“)
- Inhomogene Neumann-Randbedingung: $a\nabla u \cdot n = 1$ auf $\Gamma_{N1} := \partial\Omega \setminus (\Gamma_D \cup \Gamma_{N0})$ („Einheits-Fluss“).
- Ausgabe s : Temperaturintegral über Γ_{N1}

Schwache Form dieses Problems lautet: gesucht ist $u \in X := H_{\Gamma_D}^1(\Omega) := \{u \in H^1(\Omega) \mid u|_{\Gamma_D} = 0\}$ so dass

$$\int_{\Omega} a(x)\nabla u(x) \cdot \nabla v(x) dx = \int_{\Gamma_{N1}} v(x) dx \quad \forall v \in X.$$

Übung: Diese schwache Form besitzt:

- i) eine parametrische stetige symmetrische Bilinearform b , welche gleichmäßig beschränkt, gleichmäßig koerziv und Lipschitz-stetig bzgl. μ ist
- ii) eine separierbar parametrische stetige Linearform f , welche gleichmäßig beschränkt und Lipschitz-stetig bzgl. μ ist.

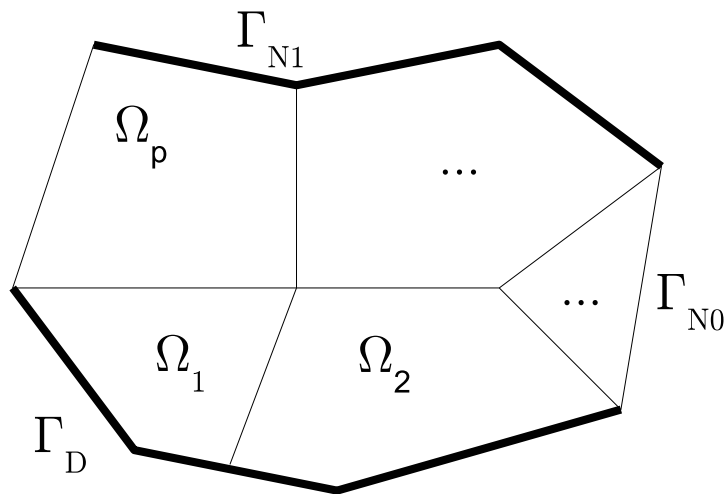


Abbildung 2.3: Komposit-Material

Kapitel 3

RB-Methode für lineare koerzive Probleme

3.1 RB-Verfahren

Definition 3.1 (Parametrisches detailliertes Problem (P)):

Sei X HR, $\mathcal{P} \subset \mathbb{R}^p$ beschränkt, $b : X \times X \times \mathcal{P} \rightarrow \mathbb{R}$ parametrische, stetige, koerzive Bilinearform, $f, l : X \times \mathcal{P} \rightarrow \mathbb{R}$ parametrische, stetige Linearform. Zu $\mu \in \mathcal{P}$ ist die detaillierte Lösung $u(\mu) \in X$ und Ausgabe $s(\mu)$ gesucht mit

$$\begin{aligned} b(u(\mu), v; \mu) &= f(v; \mu) \quad \forall v \in X \\ s(\mu) &= l(u(\mu); \mu) \end{aligned}$$

im Folgenden mit (P) bezeichnet.

Bemerkung:

- i) (P) kann sowohl ein analytisches Modell (schwache Form einer PDE) als auch ein diskretes, detailliertes Modell (FEM Diskretisierung) repräsentieren.
- ii) Existenz und Eindeutigkeit folgen mit Lax-Milgram Satz 2.11.
- iii) Beispiel 2.17 mit geeigneten Ausgabefunktional ist eine Instanz von (P) .

Definition 3.2 (Reduzierte Basis, RB-Räume):

Sei $S_N := \{\mu^1, \dots, \mu^N\} \subset \mathcal{P}$ Menge von Parametern mit (o.B.d.A) linear unabhängigen Lösungen $\{u(\mu^i)\}_{i=1}^N$ von (P) . Dann ist

$$X_N := \text{span}\{u(\mu^i)\}_{i=1}^N$$

ein N -dimensionaler Lagrange RB-Raum.

Sei $\mu^0 \in \mathcal{P}$ und $u(\mu)$ Lösung von (P) k -mal diff'bar in einer Umgebung von μ^0 . Dann ist

$$X_{k,\mu^0} := \text{span}\{\partial_\sigma u(\mu^0) \mid \sigma \in \mathbb{N}_0^p, |\sigma| \leq k\}$$

ein Taylor RB-Raum. Eine Basis

$$\Phi_N = \{\varphi_1, \dots, \varphi_N\} \subset X$$

eines RB-Raumes ist eine reduzierte Basis.

Bemerkung:

- i) Es existieren weitere Arten von RB-Räumen (Hermite, POD). Gemeinsam ist allen die „Konstruktion aus Snapshots von u und $\partial_\sigma u$ “
- ii) Für numerische Stabilität meist Verwendung von orthonormalen Basen
 \Rightarrow Details in Kapitel 5

Definition 3.3 (RB-Modell (P_N) , symmetrischer Fall):

Sei ein Problem (P) gegeben und zusätzlich gelte b symmetrisch und $f = l$ („compliant“). Sei $X_N \subset X$ ein RB-Raum. Zu $\mu \in \mathcal{P}$ ist die RB-Lösung $u_N(\mu) \in X_N$ und die RB-Ausgabe $s_N(\mu) \in \mathbb{R}$ gesucht mit

$$\begin{aligned} b(u_N(\mu), v; \mu) &= f(v; \mu) \quad \forall v \in X_N \\ s_N(\mu) &= l(u_N(\mu); \mu) \end{aligned}$$

im Folgenden mit (P_N) bezeichnet.

Bemerkung:

Falls b nicht symmetrisch oder $f \neq l$ ist obiges immer noch sinnvoll, aber es bestehen bessere Möglichkeiten $s_N(\mu)$ mittels eines dualen Problems zu bestimmen, siehe Kapitel 7.

Satz 3.4 (Existenz, Eindeutigkeit und Beschränktheit):

Zu $\mu \in \mathcal{P}$ existiert eine eindeutige RB-Lösung $u_N(\mu) \in X_N$ und RB-Ausgabe $s_N(\mu)$ von (P_N) . Diese sind beschränkt durch

$$\begin{aligned} \|u_N(\mu)\| &\leq \frac{1}{\alpha(\mu)} \|f(\cdot; \mu)\| \\ |s_N(\mu)| &\leq \frac{1}{\alpha(\mu)} \|f(\cdot; \mu)\| \|l(\cdot; \mu)\| \end{aligned}$$

Beweis: Da $X_N \subset X$ ist $b(\cdot, \cdot; \mu)$ stetig und koerziv auf X_N :

$$\alpha_N(\mu) := \inf_{u \in X_N} \frac{b(u, u; \mu)}{\|u\|^2} \geq \inf_{u \in X} \frac{b(u, u; \mu)}{\|u\|^2} = \alpha(\mu) > 0,$$

$$\gamma_N(\mu) := \sup_{u, v \in X_N} \frac{b(u, v; \mu)}{\|u\| \|v\|} \leq \sup_{u, v \in X} \frac{b(u, v; \mu)}{\|u\| \|v\|} = \gamma(\mu) < \infty.$$

Analog sind f, l stetig auf X_N . Existenz und Eindeutigkeit folgt mit Lax-Milgam, Satz 2.11.

$$\begin{aligned} b(u_N, v; \mu) &= f(v; \mu) \quad \forall v \in X_N \\ \stackrel{2.11}{\Rightarrow} (Bu_N, v) &= (v_f, v) \quad \forall v \in X_N \\ \Rightarrow Bu_N &= v_f \\ \Rightarrow u_N &= B^{-1}v_f \end{aligned}$$

Damit ist auch $s_n(\mu) = l(u_N(\mu); \mu)$ eindeutig. Beschränktheit:

$$\|u_N\| = \|B^{-1}v_f\| \leq \underbrace{\|B^{-1}\|}_{\leq \frac{1}{\alpha(\mu)}} \underbrace{\|v_f\|}_{=\|f\|}$$

$$|s_N(\mu)| = |l(u_N)| \leq \|l\| \|u_N\| \leq \frac{1}{\alpha(\mu)} \|f\| \|l\|.$$

□

Satz 3.5 (Lipschitz-Stetigkeit):

Falls b und f gleichmäßig beschränkt und Lipschitz-stetig und b gleichmäßig koerziv bzgl. μ , so sind auch die Lösungen $u_N(\mu)$ und $s_N(\mu)$ von (P_N) Lipschitz-stetig bzgl. μ .

Beweis: Siehe Übung.

□

Satz 3.6 (Gleichungssystem und numerische Stabilität):

Sei $\Phi_N = \{\varphi_1, \dots, \varphi_N\}$ eine reduzierte Basis von X_N . Für $\mu \in \mathcal{P}$ definieren wir

$$\begin{aligned} \underline{B}_N(\mu) &:= (b(\varphi_i, \varphi_j; \mu))_{i,j=1}^N \in \mathbb{R}^{N \times N} \\ \underline{f}_N(\mu) &:= \underline{f}_N(\mu) := (f(\varphi_j; \mu))_{j=1}^N \in \mathbb{R}^N \end{aligned}$$

und $\underline{u}_N = (u_{N,i})_{i=1}^N \in \mathbb{R}^N$ als Lösung von

$$\underline{B}_N(\mu) \cdot \underline{u}_N = \underline{f}_N(\mu). \quad (3.1)$$

- i) Dann ist $u_N(\mu) = \sum u_{N,i}\varphi_i$ und $s_N(\mu) = \underline{l}_N^T \cdot \underline{u}_N(\mu)$ Lösung von (P_N) .
- ii) Falls Φ_N orthonormal, so ist die Kondition von (3.1) unabhängig von N beschränkt durch

$$\text{cond}_2(\underline{B}_N(\mu)) \stackrel{\text{Def}}{=} \|\underline{B}_N(\mu)\|_2 \cdot \|\underline{B}_N^{-1}(\mu)\|_2 \leq \frac{\gamma(\mu)}{\alpha(\mu)}.$$

Beweis:

- i) folgt durch Einsetzen.
- ii) Wegen Symmetrie ist

$$\text{cond}_2(\underline{B}_N(\mu)) = \frac{|\lambda_{max}|}{|\lambda_{min}|} \quad (3.2)$$

mit betragsmäßig größtem/kleinsten Eigenwert λ_{max} , λ_{min} von $\underline{B}_N(\mu)$. Sei $\underline{u}_{max} = (u_i)_{i=1}^N \in \mathbb{R}^N$ Eigenvektor zu λ_{max} und $u_{max} := \sum_{i=1}^N u_i \varphi_i$. Dann gilt

$$\begin{aligned} \lambda_{max} \|\underline{u}_{max}\|^2 &= \lambda_{max} \underline{u}_{max}^T \underline{u}_{max} = \underline{u}_{max}^T \cdot \underline{B}_N(\mu) \cdot \underline{u}_{max} \\ &= \sum_{i,j} u_i u_j b(\varphi_i, \varphi_j; \mu) \\ &= b\left(\sum_i u_i \varphi_i, \sum_j u_j \varphi_j; \mu\right) \\ &\leq \gamma(\mu) \|\underline{u}_{max}\|^2. \end{aligned}$$

Mit

$$\|\underline{u}_{max}\|^2 = \left(\sum_{i=1}^N u_i \varphi_i, \sum_{j=1}^N u_j \varphi_j \right) = \sum_{i,j=1}^N u_i u_j (\varphi_i, \varphi_j) = \sum_{i=1}^N u_i^2 = \|\underline{u}_{max}\|^2$$

folgt $|\lambda_{max}| \leq \gamma(\mu)$. Analog zeigt man $\lambda_{min} \geq \alpha(\mu)$ und mit (3.2) folgt die Behauptung. □

Bemerkung (Unterschied FEM zu RB):

- i) Die Kondition der Matrix bei FEM wächst polynomiell in Anzahl der Freiheitsgrade. Obige Aussage ist also ein konzeptioneller Vorteil von RB-Verfahren, wenn eine orthonormale Basis Φ_N verwendet wird.

- ii) im Gegensatz zu großen, aber dünn besetzten Matrizen bei FEM, ist (3.1) klein (typisch (10-100)) aber voll besetzt, weil φ_i i.a. keine disjunkte Träger haben.

Korollar 3.7 (Galerkin-Projektion, Galerkin-Orthogonalität):

Sei $\mu \in \mathcal{P}$, X, X_N HR mit Energie-Skalarprodukt $((\cdot, \cdot))_\mu$, $\mathcal{P}_\mu : X \rightarrow X_N$ zugehörige orthogonale Projektion, $u(\mu), u_N(\mu)$ Lösungen von (P) bzw. (P_N) und der Fehler $e(\mu) := u(\mu) - u_N(\mu)$. Dann gilt

i) $u_N(\mu) = \mathcal{P}_\mu(u(\mu))$ „Galerkin-Projektion“

ii) $((e(\mu), v))_\mu = 0, v \in X_N$ „Galerkin-Orthogonalität“

Beweis: Nach Lemma 2.16 ist $(X, ((\cdot, \cdot))_\mu)$ HR und $X_N = \text{span}\{\varphi_i\}_{i=1}^N$ endlich dimensional, also abgeschlossen, also ist \mathcal{P}_μ nach 2.3 wohldefiniert. Orthogonalität des Projektionsfehlers Satz 2.3 i) ergibt

$$\begin{aligned} & ((\mathcal{P}_\mu u(\mu) - u(\mu), \varphi_i))_\mu = 0 \quad \forall i \\ \Leftrightarrow & b(\mathcal{P}_\mu u(\mu) - u(\mu), \varphi_i; \mu) = 0 \quad \forall i \\ & \Leftrightarrow b(\mathcal{P}_\mu u(\mu), \varphi_i; \mu) = b(u(\mu), \varphi_i; \mu) \quad \forall i \\ & \Leftrightarrow b(\mathcal{P}_\mu u(\mu), \varphi_i; \mu) = f(\varphi_i; \mu) \quad \forall i \end{aligned}$$

i) Also ist $\mathcal{P}_\mu u(\mu)$ Lösung von (P_N) , also $u_N(\mu) = \mathcal{P}_\mu u(\mu)$.

ii) $e(\mu)$ ist also Projektionsfehler, orthogonal nach Satz 2.3 i).

□

Korollar 3.8 (Reproduktion von Lösungen):

Sei $\mu \in \mathcal{P}$, $u(\mu), u_N(\mu)$ Lösungen von (P) bzw. (P_N) , $\underline{e}_i \in \mathbb{R}^N$ i -ter Einheitsvektor.

i) Falls $u(\mu) \in X_N \Rightarrow u_N(\mu) = u(\mu)$

ii) Falls $u(\mu) = \varphi_i \in \Phi_N$, dann $\underline{u}_N(\mu) = \underline{e}_i \in \mathbb{R}^N$

Beweis:

i) Da $u(\mu), u_N(\mu) \in X_N \Rightarrow e := u(\mu) - u_N(\mu) \in X_N$ und $((e, v))_\mu = 0 \quad \forall v \in X_N$ nach Korollar 3.7 ii), also insbesondere $((e, e))_\mu = 0 \Rightarrow e = 0 \in X_N \Rightarrow u_N(\mu) = u(\mu)$.

ii) $u_N(\mu) = \varphi_i$ nach i), mit Eindeutigkeit der Basisexpansion folgt die Behauptung.

□

3.2 Offline/Online-Zerlegung des RB-Modells

Komplexitätsbetrachtungen

Falls (P) ein diskretes Modell mit $\dim X = H$ Freiheitsgraden und das $H \times H$ LGS dünn besetzt ist, erfordert eine detaillierte Lösung $O(H^2)$ Rechenschritte. Das voll besetzte reduzierte $N \times N$ System aus (P_N) ist in $O(N^3)$ lösbar \Rightarrow Nur falls $N \ll H$ ist das ein Gewinn.

Genauere Betrachtung der Berechnung einer reduzierten Lösung $u_N(\mu)$

- i) N Snapshots berechnen mittels (P) : $O(NH^2)$
- ii) N^2 Auswertungen von $b(\varphi_i, \varphi_j; \mu)$: $O(N^2H)$
- iii) N Auswertungen von $f(\varphi_i, \mu)$: $O(NH)$
- iv) Lösen des $N \times N$ LGS: $O(N^3)$

Für ein einzelnes $\mu \in \mathcal{P}$ lohnt sich ein RB-Modell nicht. Für viele Parameter lohnt es sich durch eine Offline/Online-Zerlegung.

In der einmalig durchgeführten Offline-Phase werden μ -unabhängige Größen typischerweise teuer berechnet, d.h. $O(H^\alpha)$, $\alpha \in \mathbb{N}$. In der vielfach durchgeführten Online-Phase werden μ -abhängige Größen assembliert und $u_N(\mu)$, $s_N(\mu)$ schnell berechnet, d.h. $O(N^\alpha)$, $\alpha \in \mathbb{N}$.

Korollar 3.9 (Offline/Online-Zerlegung von (P_N)):

Sei (P_N) gegeben und b, f separierbar parametrisch, d.h.

$$b(u, v; \mu) = \sum_{q=1}^{Q_b} \Theta_b^q(\mu) b^q(u, v), \quad f(v; \mu) = \sum_{q=1}^{Q_f} \Theta_f^q(\mu) f^q(v)$$

(Offline-Phase:) Nach Berechnung einer reduzierten Basis $\Phi_N = \{\varphi_1, \dots, \varphi_N\}$ ermitteln wir

$$\begin{aligned} \underline{B}_N^q &:= (B_{N,i,j}^q)_{i,j=1}^N \in \mathbb{R}^{N \times N} & B_{N,i,j}^q &:= b^q(\varphi_i, \varphi_j) \\ \underline{f}_N^q &:= (f_{N,i}^q)_{i=1}^N \in \mathbb{R}^N & f_{N,i}^q &:= f^q(\varphi_i) \end{aligned}$$

(Online-Phase:) Zu $\mu \in \mathcal{P}$ erhalten wir $\underline{B}_N(\mu), \underline{f}_N(\mu)$ aus Satz 3.6 äquivalent

durch

$$\underline{B}_N(\mu) = \sum_{q=1}^{Q_b} \Theta_b^q(\mu) \cdot \underline{B}_N^q \quad \text{und}$$

$$\underline{f}_N(\mu) = \sum_{q=1}^{Q_f} \Theta_f^q(\mu) \cdot \underline{f}_N^q$$

und erhalten $u_N(\mu)$ und $s_N(\mu)$ wie in Satz 3.6 Insbesondere ist die Online-Phase in $O(N^2 \cdot Q_b + N \cdot Q_f + N^3)$ berechenbar, vollkommen unabhängig von H .

Diagramm: Laufzeit für $k \in \mathbb{N}$ Berechnungen von $u(\mu), u_N(\mu)$

Es bezeichnet $t_{detail}, t_{offline}, t_{online}$ Laufzeiten für einzelne Berechnungen von (P) , Offline-Phase von (P_N) und Online-Phase von (P_N) .

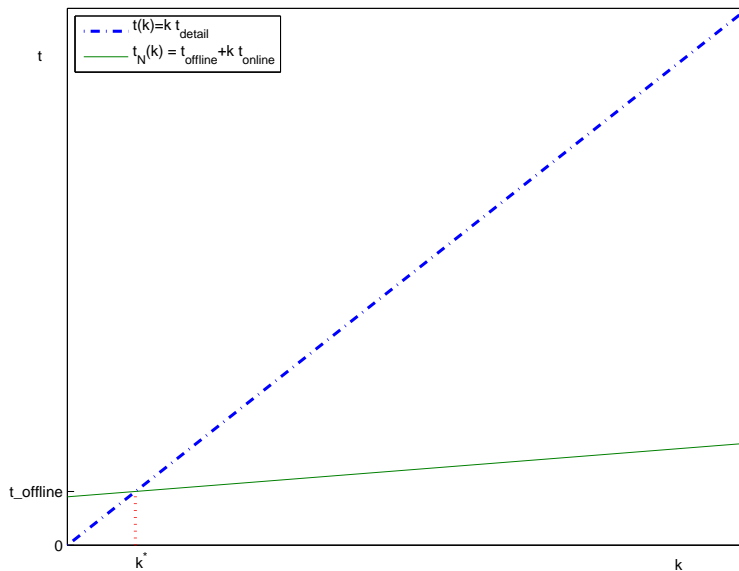


Abbildung 3.1: Laufzeitverhalten mit wachsender Anzahl von Simulationen

\Rightarrow Bei mehr als $k^* := \frac{t_{offline}}{t_{detail} - t_{online}}$ wiederholten Berechnungen von $u_N(\mu)$ lohnt sich ein RB-Modell.

3.3 Fehleranalyse

Satz 3.10 (Fehlerschranke durch Approximationsfehler):

Sei $\mu \in \mathcal{P}$ und $u(\mu), s(\mu)$ bzw. $u_N(\mu), s_N(\mu)$ Lösung von (P) bzw. (P_N)

i) Der Fehler in der (μ -abhängigen) Energienorm erfüllt

$$\|u(\mu) - u_N(\mu)\|_\mu = \inf_{v \in X_N} \|u(\mu) - v\|_\mu$$

ii) Der Fehler in der (μ -unabhängigen) X -Norm erfüllt

$$\|u(\mu) - u_N(\mu)\| \leq \sqrt{\frac{\gamma(\mu)}{\alpha(\mu)}} \inf_{v \in X_N} \|u(\mu) - v\|$$

mit $\gamma(\mu), \alpha(\mu)$ Stetigkeits- bzw. Koerzivitätskonstante.

iii) Für den Ausgabefehler gilt (wegen $l = f$)

$$\begin{aligned} 0 \leq s(\mu) - s_N(\mu) &= \|u(\mu) - u_N(\mu)\|_\mu^2 \\ &= \inf_{v \in X_N} \|u(\mu) - v\|_\mu^2 \\ &\leq \gamma(\mu) \inf_{v \in X_N} \|u(\mu) - v\|^2 \end{aligned}$$

Beweis:

i) Nach Kor. 3.7 ist $u_N(\mu)$ orthogonale Projektion also Bestapproximation:

$$\|u(\mu) - u_N(\mu)\|_\mu \stackrel{3.7i)}{=} \|u(\mu) - \mathcal{P}_\mu(u(\mu))\|_\mu \stackrel{2.3ii)}{=} \inf_{v \in X_N} \|u(\mu) - v\|_\mu$$

ii) Mit der Normäquivalenz 2.16 folgt

$$\begin{aligned} \sqrt{\alpha(\mu)} \|u(\mu) - u_N(\mu)\| &\stackrel{2.16}{\leq} \|u(\mu) - u_N(\mu)\|_\mu \stackrel{i)}{=} \inf_{v \in X_N} \|u(\mu) - v\|_\mu \\ &\stackrel{2.16}{\leq} \sqrt{\gamma(\mu)} \inf_{v \in X_N} \|u(\mu) - v\| \end{aligned}$$

iii)

$$\begin{aligned} s(\mu) - s_N(\mu) &\stackrel{Def}{=} l(u(\mu)) - l(u_N(\mu)) \\ &\stackrel{l=f}{=} f(u(\mu)) - f(u_N(\mu)) \\ &\stackrel{(P), 3.7}{=} b(u(\mu), u(\mu) - u_N(\mu)) - \underbrace{b(u_N(\mu), u(\mu) - u_N(\mu))}_{=0 \text{ wg. Kor. 3.7}} \\ &= b(u - u_N, u - u_N) \stackrel{i)}{=} \|u(\mu) - u_N(\mu)\|_\mu^2 \\ &= \inf_{v \in X_N} \|u(\mu) - v\|_\mu^2 \\ &\stackrel{2.16}{\leq} \gamma(\mu) \cdot \inf_{v \in X_N} \|u(\mu) - v\|^2 \end{aligned}$$

Also insbesondere auch $s(\mu) - s_N(\mu) \geq 0$

□

Bemerkung:

- i) $s_N(\mu)$ ist also untere Schranke für $s(\mu)$
- ii) Der Ausgabefehler ist im Allgemeinen sehr klein, da das Quadrat des RB-Fehlers eingeht
- iii) „Lemma von Cea“ besagt für nicht notwendigerweise symmetrische Bilinearformen

$$\|u - u_N\| \leq \frac{\gamma}{\alpha} \inf_{v \in X_N} \|u - v\|$$

Also ist Satz 3.10 eine Verschärfung von Cea für symmetrische Bilinearformen.

Korollar 3.11 (Monotoner Fehlerabfall in der Energienorm):

Sei $(X_N)_{N=1}^{N_{max}}$ Folge von RB-Räumen mit $X_N \subseteq X_{N'}$ für $N \leq N'$ („Hierarchische Räume“) und $e_N(\mu) = u(\mu) - u_N(\mu)$ für $\mu \in \mathcal{P}$. Dann ist die Folge $(\|e_N(\mu)\|_\mu)_{N=1}^{N_{max}}$ monoton fallend.

Beweis:

$$\begin{aligned} \|e_N(\mu)\|_\mu &= \inf_{v \in X_N} \|u(\mu) - v\|_\mu \geq \inf_{v \in X_{N'}} \|u(\mu) - v\|_\mu \\ &= \|e_{N'}(\mu)\|_\mu \end{aligned}$$

□

Bemerkung:

- i) „worst case“ ist Stagnation des Fehlers (unrealistisch, jeder neue Basisvektor müsste orthogonal zu $e_N(\mu)$ sein), siehe Abb. 3.3 a)
in der Praxis ist bei geschickter Basiswahl exponentielle Konvergenz zu beobachten, siehe Kapitel 4.
- ii) Monotonie gilt nicht notwendigerweise für andere Normen trotz Normenäquivalenz $c\|e_N\|_\mu \leq \|e_N\| \leq C\|e_N\|_\mu$ mit c, C unabhängig von

N , siehe Abb. 3.3 b). Fehlernorm $\|e_N\|$ kann gelegentlich anwachsen, bleibt aber in einem „Korridor“ um $\|e_N(\mu)\|_\mu$

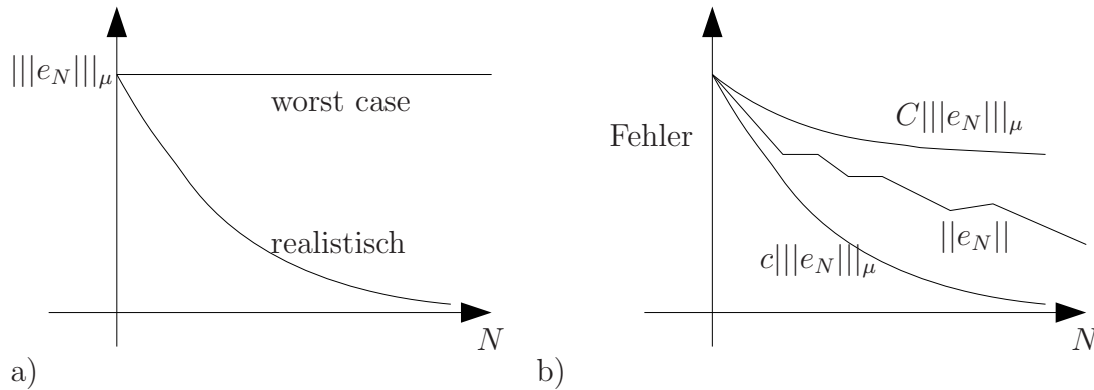


Abbildung 3.2: Fehlerabfall mit wachsender reduzierter Dimension.

Lemma 3.12 (Fehler-Residuum-Beziehung):

Für $\mu \in \mathcal{P}$ definieren wir mittels der RB-Lösung $u_N(\mu)$ das Residuum $r(\cdot; \mu) \in X'$ bzw. seinen Riesz-Repräsentanten $v_r(\mu) \in X$

$$(v_r(\mu), v) := r(v; \mu) := f(v; \mu) - b(u_N(\mu), v; \mu) \quad \forall v \in X$$

Dann erfüllt der Fehler $e(\mu) = u(\mu) - u_N(\mu)$

$$b(e(\mu), v; \mu) = r(v; \mu) \quad \forall v \in X$$

Beweis:

$$\begin{aligned} b(e, v; \mu) &\stackrel{Def}{=} b(u - u_N, v) = b(u, v) - b(u_N, v) \\ &= f(v) - b(u_N, v) = r(v) \end{aligned}$$

□

Bemerkung:

- i) Der Fehler erfüllt also (P) mit „rechter Seite“ r .
- ii) Insbesondere ist $r(v) = 0$ für $v \in X_N$ (Galerkin-Orth.)
- iii) $r = 0 \Rightarrow e = 0$.

Satz 3.13 (A-posteriori Fehlerschätzer):

Sei $\mu \in \mathcal{P}$, $u(\mu), s(\mu)$ Lösung von (P) , $u_N(\mu), s_N(\mu)$ Lösung von (P_N) für

$f \neq 0$ und der Fehler $e(\mu) := u(\mu) - u_N(\mu)$. Sei $\alpha_{LB}(\mu) > 0$ eine untere Schranke für die Koerzivitätskonstante $\alpha(\mu)$ von $b(\cdot, \cdot; \mu)$ und $v_r(\mu)$ der Riesz-Repräsentant des Residuums aus Lemma 3.12 auf. Dann gelten folgende Schranken:

i) Für den Fehler in der Energienorm:

$$\| \|e(\mu)\| \|_{\mu} \leq \Delta_N^{en}(\mu) := \frac{\|v_r(\mu)\|}{\sqrt{\alpha_{LB}(\mu)}}$$

ii) Für den relativen Fehler in der Energienorm:

$$\frac{\| \|e(\mu)\| \|_{\mu}}{\| \|u(\mu)\| \|_{\mu}} \leq \Delta_N^{en,rel}(\mu) := 2 \cdot \frac{\|v_r(\mu)\|}{\sqrt{\alpha_{LB}(\mu)}} \cdot \frac{1}{\| \|u_N(\mu)\| \|_{\mu}},$$

falls $\Delta_N^{en,rel}(\mu) \leq 1$

iii) Für den Fehler in $\|\cdot\|$ (X -Norm)

$$\|e(\mu)\| \leq \Delta_N(\mu) := \frac{\|v_r(\mu)\|}{\alpha_{LB}(\mu)}$$

iv) Für den relativen Fehler in $\|\cdot\|$

$$\frac{\|e(\mu)\|}{\|u(\mu)\|} \leq \Delta_N^{rel}(\mu) := 2 \cdot \frac{\|v_r(\mu)\|}{\alpha_{LB}(\mu)} \cdot \frac{1}{\|u_N(\mu)\|}, \text{ falls } \Delta_N^{rel} \leq 1$$

v) Für den Ausgabefehler

$$s(\mu) - s_N(\mu) \leq \Delta_N^s(\mu) := \frac{\|v_r(\mu)\|^2}{\alpha_{LB}(\mu)}$$

vi) Für den relativen Ausgabefehler

$$\frac{s(\mu) - s_N(\mu)}{s(\mu)} \leq \Delta_N^{s,rel}(\mu) := \frac{\|v_r(\mu)\|^2}{\alpha_{LB}(\mu)s_N(\mu)}$$

Bemerkung:

i) Da $f \neq 0$ folgt $u, u_N \neq 0$ und damit $s, s_N \neq 0$:

$$s(\mu) = l(u(\mu); \mu) = f(u(\mu); \mu) = b(u(\mu), u(\mu); \mu) \geq \alpha(\mu)\|u(\mu)\|^2 > 0,$$

also alle Brüche in i)–vi) sinnvoll.

- ii) X -Norm Fehlerschätzer immer um Faktor $\sqrt{\alpha_{LB}(\mu)}$ schlechter als Energienorm-Fehlerschätzer
- iii) sehr gute Ausgabefehlerschätzer durch „Quadrat“ des Residuums

Beweis: Zunächst gilt

$$\|e\|_{\mu}^2 \stackrel{Def}{=} b(e, e; \mu) \stackrel{3.12}{=} r(e; \mu) \stackrel{Riesz}{=} (v_r, e) \stackrel{CSU}{\leq} \|v_r\| \|e\| \quad (3.3)$$

Andererseits folgt aus der Koerzivität

$$\alpha_{LB}(\mu) \|e\|^2 \leq b(e, e; \mu) = \|e(\mu)\|_{\mu}^2$$

also

$$\|e\| \leq \frac{\|e\|_{\mu}}{\sqrt{\alpha_{LB}(\mu)}} \quad (3.4)$$

i)

$$\|e\|_{\mu}^2 \stackrel{(3.3)}{\leq} \|v_r\| \|e\| \stackrel{3.4}{\leq} \|v_r\| \frac{\|e\|_{\mu}}{\sqrt{\alpha_{LB}(\mu)}} = \Delta_N^{en}(\mu) \|e\|_{\mu}$$

ii) Falls $\Delta_N^{en,rel}(\mu) \leq 1$, so ist

$$\begin{aligned} \left| \frac{\|u\|_{\mu} - \|u_N\|_{\mu}}{\|u_N\|_{\mu}} \right| &\stackrel{\Delta-Ungl.}{\leq} \frac{\|u - u_N\|_{\mu}}{\|u_N\|_{\mu}} = \frac{\|e\|_{\mu}}{\|u_N\|_{\mu}} \\ &\stackrel{i)}{\leq} \frac{\|v_r\|}{\sqrt{\alpha_{LB}(\mu)} \|u_N\|_{\mu}} \stackrel{Def}{=} \frac{1}{2} \Delta_N^{en,rel} \leq \frac{1}{2} \end{aligned}$$

Falls $\|u_N\|_{\mu} > \|u\|_{\mu}$ gilt $\|u_N\|_{\mu} - \|u\|_{\mu} \leq \frac{1}{2} \|u_N\|_{\mu}$ also

$$\frac{1}{2} \|u_N\|_{\mu} \leq \|u\|_{\mu} \quad (3.5)$$

Falls $\|u_N\|_{\mu} \leq \|u\|_{\mu}$, so ist (3.5) schon klar. Damit folgt

$$\frac{\|e\|_{\mu}}{\|u\|_{\mu}} \stackrel{i)}{\leq} \frac{\|v_r\|}{\sqrt{\alpha_{LB}(\mu)}} \cdot \frac{1}{\|u\|_{\mu}} \stackrel{(3.5)}{\leq} \frac{\|v_r\|}{\sqrt{\alpha_{LB}(\mu)}} \frac{1}{\|u_N\|_{\mu}} \cdot 2 = \Delta_N^{en,rel}(\mu)$$

iii) analog zu i)

iv) analog zu ii)

v) aus Satz 3.10 folgt

$$s(\mu) - s_N(\mu) \stackrel{3.10}{=} \|e(\mu)\|_{\mu}^2 \stackrel{i)}{\leq} (\Delta_N^{en}(\mu))^2 \stackrel{Def}{=} \Delta_N^s(\mu)$$

vi) Aus Satz 3.10 folgt $s(\mu) \geq s_N(\mu)$ also aus v)

$$\frac{s(\mu) - s_N(\mu)}{s(\mu)} \stackrel{v)}{\leq} \frac{\Delta_N^s(\mu)}{s(\mu)} \leq \frac{\Delta_N^s(\mu)}{s_N(\mu)} = \Delta_N^{s,rel}(\mu)$$

□

Satz 3.14 (Effektivitäten der Fehlerschätzer):

Mit den Voraussetzungen und Notationen aus 3.13 gelten die folgenden Abschätzungen für Effektivitäten, wobei $\gamma_{UB}(\mu)$ eine obere Schranke für $\gamma(\mu)$ ist:

i)

$$\eta_N^{en}(\mu) := \frac{\Delta_N^{en}(\mu)}{\|e(\mu)\|_\mu} \leq \sqrt{\frac{\gamma_{UB}(\mu)}{\alpha_{LB}(\mu)}}$$

ii)

$$\eta_N^{en,rel}(\mu) := \frac{\Delta_N^{en,rel}(\mu)}{\|e(\mu)\|_\mu / \|u(\mu)\|_\mu} \leq 3 \cdot \sqrt{\frac{\gamma_{UB}(\mu)}{\alpha_{LB}(\mu)}}, \text{ falls } \Delta_N^{en,rel}(\mu) \leq 1$$

iii)

$$\eta_N(\mu) := \frac{\Delta_N(\mu)}{\|e(\mu)\|} \leq \frac{\gamma_{UB}(\mu)}{\alpha_{LB}(\mu)}$$

iv)

$$\eta_N^{rel}(\mu) := \frac{\Delta_N^{rel}(\mu)}{\|e(\mu)\| / \|u\|} \leq 3 \cdot \frac{\gamma_{UB}(\mu)}{\alpha_{LB}(\mu)}, \text{ falls } \Delta_N^{rel}(\mu) \leq 1$$

v)

$$\eta_N^s(\mu) := \frac{\Delta_N^s(\mu)}{s(\mu) - s_N(\mu)} \leq \frac{\gamma_{UB}(\mu)}{\alpha_{LB}(\mu)}$$

vi)

$$\eta_N^{s,rel}(\mu) := \frac{\Delta_N^{s,rel}(\mu)}{(s(\mu) - s_N(\mu))/s(\mu)} \leq 2 \cdot \frac{\gamma_{UB}(\mu)}{\alpha_{LB}(\mu)}, \text{ falls } \Delta_N^{s,rel} \leq 1$$

Bemerkung:

Satz 3.13 besagt also, dass $\eta_N(\mu) \geq 1$, $\eta_N^{en}(\mu) \geq 1$, etc.

Beweis: Zunächst gilt mit Definition des Residuums aus Lem. 3.12

$$\|v_r\|^2 = (v_r, v_r) \stackrel{Def}{=} b(e, v; \mu) \stackrel{CSU}{\leq} \|e\|_\mu \|v_r\|_\mu.$$

Mit Normäquivalenz gilt

$$\|v_r\|_\mu \leq \sqrt{\gamma_{UB}(\mu)} \cdot \|v_r\|$$

also

$$\|v_r\| \leq \|e\|_\mu \cdot \sqrt{\gamma_{UB}(\mu)} \quad (3.6)$$

i)

$$\frac{\Delta_N^{en}(\mu)}{\|e(\mu)\|_\mu} \stackrel{Def}{=} \frac{\|v_r\|}{\|e(\mu)\|_\mu} \cdot \frac{1}{\sqrt{\alpha_{LB}(\mu)}} \stackrel{(3.6)}{\leq} \frac{\|e\|_\mu \sqrt{\gamma_{UB}(\mu)}}{\|e\|_\mu \sqrt{\alpha_{LB}(\mu)}}$$

v) Satz 3.10 iii) liefert

$$\eta_N^s(\mu) = \frac{\Delta_N^s(\mu)}{s(\mu) - s_N(\mu)} \stackrel{Def, 3.10}{=} \frac{(\Delta_N^{en}(\mu))^2}{\|e(\mu)\|_\mu^2} \stackrel{i)}{\leq} \frac{\gamma_{UB}(\mu)}{\alpha_{LB}(\mu)}$$

vi)

$$\eta_N^{s,rel} = \frac{\Delta_N^s(\mu)/s_N(\mu)}{(s(\mu) - s_N(\mu))/s(\mu)} = \eta_N^s(\mu) \cdot \frac{s(\mu)}{s_N(\mu)} \stackrel{v)}{\leq} \frac{\gamma_{UB}(\mu)}{\alpha_{LB}(\mu)} \cdot \frac{s(\mu)}{s_N(\mu)}$$

Der letzte Faktor ist beschränkt

$$\frac{s(\mu)}{s_N(\mu)} = 1 + \frac{s(\mu) - s_N(\mu)}{s_N(\mu)} \stackrel{Def}{\leq} 1 + \Delta_N^{s,rel}(\mu) \stackrel{Voraus}{\leq} 2.$$

iii) Gl. (3.6) und Normäquivalenz liefern

$$\|v_r\| \stackrel{(3.6)}{\leq} \|e\|_\mu \sqrt{\gamma_{UB}(\mu)} \leq \gamma_{UB}(\mu) \|e\|$$

Also

$$\eta_N(\mu) = \frac{\Delta_N(\mu)}{\|e(\mu)\|} \stackrel{Def}{=} \frac{\|v_r\|}{\alpha_{LB}(\mu)} \cdot \frac{1}{\|e\|} \leq \frac{\gamma_{UB}(\mu) \cdot \|e\|}{\alpha_{LB}(\mu) \cdot \|e\|}$$

ii) Wie im Beweis von 3.13 impliziert $\Delta_N^{en,rel}(\mu) \leq 1$ dass

$$\left| \frac{\|u\|_\mu - \|u_N\|_\mu}{\|u_N\|_\mu} \right| \leq \frac{1}{2}$$

Falls $|||u_N|||_\mu \leq |||u|||_\mu$, so gilt $|||u|||_\mu - |||u_N||| \leq \frac{1}{2}|||u_N|||_\mu$ also

$$|||u|||_\mu \leq \frac{3}{2}|||u_N|||_\mu. \quad (3.7)$$

Falls $|||u_N|||_\mu > |||u|||_\mu$ ist (3.7) schon klar. Also

$$\begin{aligned} \eta_N^{en,rel}(\mu) &\stackrel{Def}{=} \frac{2 \cdot |||v_r|||}{\underbrace{\sqrt{\alpha_{LB}(\mu)}|||u_N|||_\mu}_{\Delta_N^{en,rel}}} \cdot \frac{1}{|||e|||_\mu / |||u|||_\mu} \\ &\stackrel{(3.6)}{\leq} \frac{2 \cdot \sqrt{\gamma_{UB}(\mu)}|||e|||_\mu}{\sqrt{\alpha_{LB}(\mu)}|||e|||_\mu} \cdot \frac{|||u|||_\mu}{|||u_N|||_\mu} \\ &\stackrel{(3.7)}{\leq} 3 \cdot \frac{\sqrt{\gamma_{UB}(\mu)}}{\sqrt{\alpha_{LB}(\mu)}} \end{aligned}$$

iv) Analog zu ii)

□

3.4 Offline/Online-Zerlegung der Fehlerschätzer

Für die a-posteriori Fehlerschätzer aus Satz 3.13 und den Effektivitäten aus Satz 3.14 wäre der Wert für $\alpha(\mu)$ bzw. $\gamma(\mu)$ wünschenswert. Dieser ist jedoch i.A. nicht verfügbar oder nur teuer berechenbar. Daher wurden als Ersatz eine untere Schranke $\alpha_{LB}(\mu)$ bzw. obere Schranke $\gamma_{UB}(\mu)$ eingeführt, welche schnell berechenbar sein sollen, damit zur Online-Fehlerschätzung einsetzbar sind.

Lemma 3.15 (Untere Schranke für $\alpha(\mu)$):

Sei $b : X \times X \times \mathcal{P} \rightarrow \mathbb{R}$ koerziv und separierbar parametrisch mit $b^q(u, u) \geq 0 \quad \forall u \in X$ und $\Theta_b^q(\mu) > 0 \quad \forall \mu \in \mathcal{P}$. Sei $\bar{\mu} \in \mathcal{P}$ fest und $\alpha(\bar{\mu})$ bekannt. Dann gilt:

$$0 < \alpha_{LB}(\mu) \leq \alpha(\mu) \quad \forall \mu \in \mathcal{P}$$

mit der unteren Schranke

$$\alpha_{LB}(\mu) := \alpha(\bar{\mu}) \cdot \min_{q=1, \dots, Q_b} \frac{\Theta_b^q(\mu)}{\Theta_b^q(\bar{\mu})}$$

Beweis: Für alle $u \in X$ gilt

$$\begin{aligned}
b(u, u; \mu) &= \sum_{q=1}^{Q_b} \Theta_b^q(\mu) b^q(u, u) = \sum_{q=1}^{Q_b} \frac{\Theta_b^q(\mu)}{\Theta_b^q(\bar{\mu})} \Theta_b^q(\bar{\mu}) b^q(u, u) \\
&\geq \sum_{q=1}^{Q_b} \underbrace{\left(\min_{q'=1, \dots, Q_b} \frac{\Theta_b^{q'}(\mu)}{\Theta_b^{q'}(\bar{\mu})} \right)}_{C(\mu)} \Theta_b^q(\bar{\mu}) b^q(u, u) = C(\mu) \cdot b(u, u; \bar{\mu}) \\
&\geq C(\mu) \cdot \alpha(\bar{\mu}) \|u\|^2 = \alpha_{LB}(\mu) \|u\|^2
\end{aligned}$$

Also insbesondere $\alpha(\mu) := \inf_{u \in X} \frac{b(u, u; \mu)}{\|u\|^2} \geq \alpha_{LB}(\mu)$. \square

Bemerkung:

Es gilt auch für nicht-symmetrisches b .

Lemma 3.16 (Obere Schranke für $\gamma(\mu)$):

Sei $b : X \times X \times \mathcal{P} \rightarrow \mathbb{R}$ symmetrisch, koerziv und separierbar parametrisch mit b^q positiv semidefinit und $\Theta_b^q(\mu) > 0 \quad \forall \mu \in \mathcal{P}, q = 1, \dots, Q_b$. Sei $\bar{\mu} \in \mathcal{P}$ fest und $\gamma(\bar{\mu})$ bekannt. Dann gilt

$$\gamma(\mu) \leq \gamma_{UB}(\mu) \quad \forall \mu \in \mathcal{P}$$

mit der oberen Schranke $\gamma_{UB}(\mu) := \gamma(\bar{\mu}) \max_{q=1, \dots, Q_b} \frac{\Theta_b^q(\mu)}{\Theta_b^q(\bar{\mu})}$.

Beweis: Übung. \square

Bemerkung:

Durch „Tauschen“ des Skalarproduktes/Norm auf X können die Fehlerschätzer und Effektivitäten verbessert werden, ohne die Ausgabe-Schätzung des reduzierten Modells oder die reduzierte Lösung zu verändern.

Wähle $\bar{\mu} \in \mathcal{P}$ fest und setze $X_{\bar{\mu}} := (X, (((\cdot, \cdot)))_{\bar{\mu}})$. Wegen Normäquivalenz, Lemma 2.16, gilt b bilinear, stetig, koerziv auf $X \Leftrightarrow b$ bilinear, stetig, koerziv auf $X_{\bar{\mu}}$. l, f linear, stetig auf $X \Leftrightarrow l, f$ linear, stetig auf $X_{\bar{\mu}}$.

Offensichtlich ist

$$\alpha(\bar{\mu}) := \inf_{u \in X_{\bar{\mu}}} \frac{b(u, u; \bar{\mu})}{\|u\|_{\bar{\mu}}^2} = 1$$

und

$$\begin{aligned}
\gamma(\bar{\mu}) &:= \sup_{u, v \in X_{\bar{\mu}}} \frac{b(u, v; \bar{\mu})}{\|u\|_{\bar{\mu}} \|v\|_{\bar{\mu}}} = \sup_{u, v} \frac{(((u, v)))_{\bar{\mu}}}{\|u\|_{\bar{\mu}} \|v\|_{\bar{\mu}}} \\
&\leq \sup_{u, v} \frac{\|u\|_{\bar{\mu}} \|v\|_{\bar{\mu}}}{\|u\|_{\bar{\mu}} \|v\|_{\bar{\mu}}} = 1
\end{aligned}$$

also Lemma 2.16: Normäquivalenz mit „Gleichheit“, $\gamma(\bar{\mu})/\alpha(\bar{\mu}) = 1$. Dies kann nicht verbessert werden, da nach Definition stets $\gamma(\mu)/\alpha(\mu) \geq 1$.

Lemma 3.17 (Separable Parameterabhängigkeit von $v_r(\mu)$):

Seien b, f separierbar parametrisch. Nach Riesz, Satz 2.4, existieren $v_f^q \in X$ mit

$$(v_f^q, v) = f^q(v) \quad \forall v \in X, 1 \leq q \leq Q_f$$

und $v_b^{q,n} \in X$ mit

$$(v_b^{q,n}, v) = b^q(\varphi_n, v) \quad \forall v \in X, 1 \leq q \leq Q_b, 1 \leq n \leq N$$

Sei $Q_r := Q_f + N \cdot Q_b$ und v_r^q für $1 \leq q \leq Q_r$ eine Aufzählung von $v_f^q, v_b^{q,n}$:

$$(v_r^1, \dots, v_r^{Q_r}) := (v_f^1, \dots, v_f^{Q_f}, v_b^{1,1}, \dots, v_b^{Q_b,1}, v_b^{1,2}, \dots, v_b^{Q_b,2}, \dots, v_b^{1,N}, \dots, v_b^{Q_b,N})$$

Für $\mu \in \mathcal{P}$ sei $u_N(\mu) = \sum_{n=1}^N u_{Nn}(\mu)\varphi_n$ die Lösung von (P_N) und hiermit definieren wir $\Theta_r^q : \mathcal{P} \rightarrow \mathbb{R}, 1 \leq q \leq Q_r$ durch

$$\begin{aligned} (\Theta_r^1, \dots, \Theta_r^{Q_r}) &= (\Theta_f^1, \dots, \Theta_f^{Q_f} - \Theta_b^1 \cdot u_{N1}, \dots, -\Theta_b^{Q_b} u_{N1}, \\ &\quad - \Theta_b^1 \cdot u_{N2}, \dots, -\Theta_b^{Q_b} u_{N2}, \dots, -\Theta_b^1 u_{NN}, \dots, -\Theta_b^{Q_b} u_{NN}) \end{aligned}$$

Dann ist der Riesz-Repräsentant $v_r(\mu) \in X$ des Residuums separierbar parametrisch

$$v_r(\mu) = \sum_{q=1}^{Q_r} \Theta_r^q(\mu) v_r^q$$

Beweis: Mit Lemma 3.12 und separabler Parameterabhängigkeit ist

$$\begin{aligned} (v_r(\mu), v) &= r(v; \mu) = f(v; \mu) - b(u_N(\mu), v; \mu) \\ &= \sum_{q=1}^{Q_f} \Theta_f^q(\mu) f^q(v) - \sum_{q=1}^{Q_b} \sum_{n=1}^N \Theta_b^q(\mu) u_{Nn}(\mu) b^q(\varphi_n, v) \\ &= \sum_{q=1}^{Q_f} \Theta_f^q(\mu) (v_f^q, v) - \sum_{q=1}^{Q_b} \sum_{n=1}^N \Theta_b^q(\mu) u_{Nn}(\mu) (v_b^{q,n}, v) \quad \forall v \in X \\ &= \underbrace{\left(\sum_{q=1}^{Q_f} \Theta_f^q(\mu) v_f^q - \sum_{q=1}^{Q_b} \sum_{n=1}^N \Theta_b^q(\mu) u_{Nn}(\mu) v_b^{q,n}, v \right)}_{= \sum_{q=1}^{Q_r} \Theta_r^q(\mu) v_r^q} \quad \forall v \in X \end{aligned}$$

□

Satz 3.18 (Offline/Online der Fehlerschätzer/Effektivitäten):

Seien b, f separierbar parametrisch und schnell berechenbare Schranken $\alpha_{LB}(\mu)$ und $\gamma_{UB}(\mu)$ verfügbar. Dann sind die Fehlerschätzer aus Satz 3.13 und Effektivitäten aus Satz 3.14 effektiv berechenbar:

Offline-Phase: Nach der Offline-Phase des RB-Modells gemäß Satz 3.9 ermitteln wir

$$\underline{K}_N := ((\varphi_i, \varphi_j))_{i,j=1}^N \in \mathbb{R}^{N \times N}$$

Nach Berechnung von $v_r^q \in X, 1 \leq q \leq Q_r$ gemäß Lemma 3.17 bestimmen wir

$$\underline{G} := ((v_r^i, v_r^j))_{i,j=1}^{Q_r} \in \mathbb{R}^{Q_r \times Q_r}$$

Online-Phase: Zu $\mu \in \mathcal{P}$ bezeichne nach Online-Phase des RB-Modells gemäß Satz 3.9 $\underline{B}_N(\mu), u_N(\mu) = \sum_{n=1}^N u_{Nn}(\mu) \varphi_n$ (und $s_N(\mu)$). Hiermit erhalten wir Bestandteile der Fehlerschranken:

- i) $|||u_N(\mu)|||_\mu = (\underline{u}_N^T \underline{B}_N(\mu) \underline{u}_N)^{1/2}$
- ii) $||u_N(\mu)|| = (\underline{u}_N^T \underline{K}_N \underline{u}_N)^{1/2}$
- iii) $||v_r(\mu)|| = (\underline{\Theta}_r^T \underline{G} \underline{\Theta}_r)^{1/2}$

mit $\underline{\Theta}_r := (\Theta_r^1(\mu), \dots, \Theta_r^{Q_r}(\mu))^T, \underline{u}_N = (u_{Nn})_{n=1}^N$.

Beweis: i) $|||u_N(\mu)|||_\mu^2 = b(u_N(\mu), u_N(\mu); \mu) = \sum_{i,j=1}^N u_{Ni} u_{Nj} b(\varphi_i, \varphi_j; \mu) = \underline{u}_N^T \underline{B}_N(\mu) \underline{u}_N$

ii) $||u(\mu)||^2 = (u_N(\mu), u_N(\mu)) = \sum_{i=1}^N u_{Ni} u_{Nj} \cdot (\varphi_i, \varphi_j) = \underline{u}_N^T \underline{K}_N \underline{u}_N$

iii) $||v_r(\mu)||^2 = (\sum_{q=1}^{Q_r} \Theta_r^q v_r^q, \sum_{q=1}^{Q_r} \Theta_r^q v_r^q) = \sum_{i,j=1}^{Q_r} \Theta_r^i(\mu) \Theta_r^j(\mu) (v_r^i, v_r^j) = \underline{\Theta}_r^T \underline{G} \underline{\Theta}_r$.

□

Zusammenfassung: (Relevanz der A-posteriori Fehlerschätzer)

- i) Rigorese obere Schranke für tatsächlichen Fehler, nicht nur „Indikatoren“ wie in FEM. (Rigorosität wird auch „Zuverlässigkeit“, „Reliability“ genannt)
- ii) Effektivität: Die Schätzer sind keine beliebig große obere Schranke, sondern der Faktor der „Überschätzung“ ist klein und berechenbar. (Effektivität wird auch „Effizienz“, „Efficiency“ genannt). Insbesondere: $e(\mu) = 0 \Rightarrow \Delta_N^*(\mu) = 0$, also „a-posteriori“ Verifikation von exakter Approximation..

- iii) Theoretische Untermauerung der i.A. empirischen Basiswahl
- iv) Effiziente Berechnung: Durch Offline/Online-Zerlegung ist neben reduzierter Simulation eine schnelle Fehleraussage möglich.
- v) weitere Einsatzmöglichkeiten: Offline zur Basisgenerierung (siehe Kapitel 5) und Online (Dimensionswahl).

Kapitel 4

A-priori Approximationstheorie

4.1 Approximierbarkeit durch lineare Unterräume

Satz 3.10 ergab einen Zusammenhang zwischen RB-Fehler und Bestapproximation, u.a. $\|u(\mu) - u_N(\mu)\| \leq c \cdot \inf_{v \in X_N} \|u(\mu) - v\|$. Daher sind Schranken für die rechte Seite interessant.

Definition 4.1 (Kolmogorov n -Weite):

Sei \mathcal{M} eine kompakte Teilmenge im HR X . Zu einem abgeschlossenen Unterraum $Y \subset X$ nennen wir

$$d(Y, \mathcal{M}) := \sup_{v \in \mathcal{M}} \inf_{w \in Y} \|v - w\| = \sup_{v \in \mathcal{M}} \|v - P_Y v\|$$

den Abstand von Y zu \mathcal{M} . Für $n \in \mathbb{N}$ nennen wir

$$d_n(\mathcal{M}) := \inf_{Y \subset X, \dim(Y)=n} d(Y, \mathcal{M})$$

die Kolmogorov n -Weite der Menge \mathcal{M} . Als Abschwächung definieren wir

$$\bar{d}_n(\mathcal{M}) := \inf_{Y \subset \text{span}(\mathcal{M}), \dim(Y)=n} d(Y, \mathcal{M}).$$

Bemerkung:

- Dies sind also Maße für die Bestapproximation durch lineare Unterräume.
- Es gilt trivialerweise

$$d_0(\mathcal{M}) = \bar{d}_0(\mathcal{M}) = \sup_{v \in \mathcal{M}} \|v\|,$$

$$d_n(\mathcal{M}) \leq \bar{d}_n(\mathcal{M}) \forall n \in \mathbb{N}$$

und falls $n_0 := \dim(\text{span}(\mathcal{M})) < \infty$, so $d_n(\mathcal{M}) = \bar{d}_n(\mathcal{M}) = 0 \forall n \geq n_0$.

- Für $\mathcal{M} := \{u(\mu) | \mu \in \mathcal{P}\} \subset X$ mit u Lsg von (P) und $u_N(\mu)$ Lsg. von (P_N) gilt also

$$\|u_N(\mu) - u(\mu)\| \leq \sqrt{\frac{\gamma(\mu)}{\alpha(\mu)}} \inf_{v \in X_N} \|u_N(\mu) - v\| \leq d_N(\mathcal{M}) \sup_{\mu \in \mathcal{P}} \sqrt{\frac{\gamma(\mu)}{\alpha(\mu)}}.$$

Falls also $d_N(\mathcal{M})$ klein und γ/α beschränkt so RB-Fehler klein.

- Präzise Werte für d_n sind selten bekannt. Für endliche Mengen oder Einheitskugeln können aber exakte Werte oder Schranken für d_n angegeben werden.
- Beispiel: $\mathcal{M} := \{v \in X | \|v\| \leq 1\}$ erfüllt $d_n(\mathcal{M}) = \bar{d}_n(\mathcal{M}) = 1$ für alle $n < \dim(X)$ und $d_n(\mathcal{M}) = \bar{d}_n(\mathcal{M}) = 0$ für $n \geq \dim(X)$.
- Beispiel: $\mathcal{M} := [-1, 1]^m \subset X := \mathbb{R}^m$ erfüllt $d_n(\mathcal{M}) = \bar{d}_n(\mathcal{M}) = \sqrt{m-n}$ für alle $n \leq m$ und $d_n(\mathcal{M}) = \bar{d}_n(\mathcal{M}) = 0$ für $n \geq m$.

4.2 Fehlerfreie Approximationen

In speziellen Fällen können RB-Räume angegeben werden, welche fehlerfrei sind

Lemma 4.2 (Optimales X_N für Komposit-Block):

Sei ein Komposit-Block aus Beispiel 2.17 gegeben mit $p = 2$, $\mathcal{P} = [\mu_{min}, \mu_{max}]^2 \subset \mathbb{R}^+$, $\Omega = (0, 1)^2$, $\Omega_1 = (0, 1) \times (0, \frac{1}{2})$, $\Omega_2 = (0, 1) \times (\frac{1}{2}, 1)$, $\Gamma_D = [0, 1] \times \{0\}$, $\Gamma_{N1} = [0, 1] \times \{1\}$, $\Gamma_{N0} := \partial\Omega \setminus (\Gamma_D \cup \Gamma_{N1})$. Dann erfüllt der Lagrange RB-Raum $X_N := \text{span}\{u((\mu_{min}, \mu_{min})^T), u((\mu_{max}, \mu_{min})^T)\}$

$$\inf_{v \in X_N} \|u(\mu) - v\| = \inf_{v \in X_N} \| \|u(\mu) - v\|_{\mu} = 0 \quad \forall \mu \in \mathcal{P}$$

Beweis: Übung □

Lemma 4.3 (Optimales X_N für parameterunabhängiges b):

Sei b nicht parameterabhängig, f separierbar parametrisch mit Q_f Komponenten und $\mu^1, \dots, \mu^{Q_f} \in \mathcal{P}$ mit $\{f(\cdot; \mu^i)\}_{i=1}^{Q_f} \subset X$ linear unabhängig. Dann erfüllt der Lagrange-RB-Raum $X_N := \text{span}\{u(\mu^i)\}_{i=1}^{Q_f}$

$$\dim X_N = Q_f \text{ und } \inf_{v \in X_N} \|u(\mu) - v\| = \inf_{v \in X_N} \| \|u(\mu) - v\|_{\mu} = 0 \quad \forall \mu \in \mathcal{P}$$

Beweis: Übung. □

4.3 Asymptotische Approximationsaussagen

Satz 4.4 (Fink und Rheinboldt 1983 [FR83]: Lokale polynomiale Konvergenz):

Sei $\mu^0 \in U \subset \mathcal{P} \subset \mathbb{R}$ und $u(\mu)$ analytisch in Umgebung U . Sei X_{k,μ^0} der Taylor-RB-Raum für $k \in \mathbb{N}$. Dann existiert ein $B_\delta(\mu^0) \subset U$ und $C > 0$, s.d.

$$\inf_{v \in X_{k,\mu^0}} \|u(\mu) - v\| \leq C \cdot |\mu - \mu^0|^{k+1} \quad \forall \mu \in B_\delta(\mu^0)$$

Beweis: Taylor-Entwicklung

$$\begin{aligned} u(\mu) &= \sum_{i=0}^{\infty} \frac{\partial^i}{\partial \mu^i} u(\mu^0) \cdot \frac{1}{i!} (\mu - \mu^0)^i \\ &= \underbrace{\sum_{i=0}^k \frac{\partial^i}{\partial \mu^i} u(\mu^0) \cdot \frac{1}{i!} (\mu - \mu^0)^i}_{v_k(\mu)} + \underbrace{\sum_{i=k+1}^{\infty} \frac{\partial^i}{\partial \mu^i} u(\mu^0) \cdot \frac{1}{i!} (\mu - \mu^0)^{i-(k+1)} (\mu - \mu^0)^{k+1}}_{w_k(\mu)} \end{aligned}$$

Sei $\delta < 1$ s.d. $B_\delta(\mu^0) \subset U$ und $C' := \sup_i \left\| \frac{\partial^i}{\partial \mu^i} u(\mu^0) \cdot \frac{1}{i!} \right\| < \infty$

Dann gilt für $\mu \in B_\delta(\mu)$:

$$\begin{aligned} \|w_k(\mu)\| &\leq \sum_{i=k+1}^{\infty} \left\| \frac{\partial^i}{\partial \mu^i} u(\mu^0) \frac{1}{i!} \right\| |\mu - \mu^0|^{i-(k+1)} \\ &\leq C' \cdot \sum_{i=k+1}^{\infty} |\mu - \mu^0|^{i-(k+1)} \leq C' \cdot \frac{1}{1 - |\mu - \mu^0|} \\ &\leq C' \frac{1}{1 - \delta} =: C \end{aligned}$$

Damit gilt wegen $v_k(\mu) \in X_{k,\mu^0}$

$$\inf_{v \in X_{k,\mu^0}} \|u(\mu) - v\| \leq \|u(\mu) - v_k(\mu)\| = \|w_k(\mu)(u - u^0)^{k+1}\| \leq C |\mu - \mu^0|^{k+1}$$

□

Satz 4.5 (Maday & et al 2002 [MPT02]: Globale Exponentielle Konvergenz):

Sei $\mathcal{P} = [\mu_{min}, \mu_{max}] \subset \mathbb{R}^+$ mit $0 < \mu_{min} < 1, \mu_{max} := 1/\mu_{min}, b(u, v; \mu) :=$

$\mu b^1(u, v) + b^2(u, v)$ und f nicht parametrisch mit $a := \ln \frac{\mu_{max}}{\mu_{min}} > \frac{1}{2e}$ gegeben und $N_0 := 1 + \lfloor 2ea + 1 \rfloor$.

Zu $N \in \mathbb{N}, N \geq 2$ seien $\mu_{min} = \mu^1 < \dots < \mu^N = \mu_{max}$ logarithmisch äquidistant, d.h.

$$\ln(\mu^{i+1}) - \ln(\mu^i) = \frac{\ln(\mu_{max}) - \ln(\mu_{min})}{N-1}$$

und $X_N := \text{span}\{u(\mu^i)\}_{i=1}^N$ der zugehörige Lagrange RB-Raum.

Dann gilt

$$\frac{|||u(\mu) - u_N(\mu)|||_\mu}{|||u(\mu)|||_\mu} \leq e^{-\frac{N-1}{N_0-1}} \quad \forall \mu \in \mathcal{P}, N \geq N_0$$

Beweis: Siehe Prop. 3D in Patera und Rozza 2006 [PR07] □

Bemerkung:

- Beispielsweise ist der Satz erfüllt mit einem beliebigen Komposit-Block aus Bsp. 2.17 aus 2 Teilgebieten, bei dem ein Wärmeleitkoeffizient auf 1 fixiert ist.
- Verallgemeinerungen für multi-Parameter Szenarien sind in Arbeit.

Korollar 4.6 (Globale Exponentielle Konvergenz des Approximationsfehlers):

Mit Satz 4.5 gilt insbesondere mit $C > 0$ unabhängig von N

$$\inf_{v \in X_N} |||u(\mu) - v||| \leq C \cdot e^{-\frac{N-1}{N_0-1}} \quad \forall \mu \in \mathcal{P}, N \geq N_0$$

Beweis: Analog zu Satz 3.4 gilt $|||u(\mu)||| \leq \frac{1}{\alpha(\mu)} ||f(\mu)||_{X'}$, also gilt für die Energienorm

$$(*) \quad |||u(\mu)|||_\mu \stackrel{2.15}{\leq} \sqrt{\gamma(\mu)} ||u(\mu)|| \leq \frac{\sqrt{\gamma(\mu)}}{\alpha(\mu)} ||f(\mu)|| \leq C'$$

mit $C' := \sup_\mu \frac{\sqrt{\gamma(\mu)}}{\alpha(\mu)} ||f(\mu)||$. Also

$$\begin{aligned} \inf_{v \in X_N} |||u(\mu) - v||| &\leq |||u(\mu) - u_N(\mu)||| \stackrel{2.15}{\leq} \frac{1}{\sqrt{\alpha(\mu)}} |||u(\mu) - u_N(\mu)|||_\mu \frac{|||u(\mu)|||}{|||u(\mu)|||_\mu} \\ &\leq \frac{|||u(\mu)|||_\mu}{\sqrt{\alpha(\mu)}} e^{-\frac{N-1}{N_0-1}} \leq \underbrace{\frac{C'}{\sqrt{\alpha(\mu)}}}_{=: C} e^{-\frac{N-1}{N_0-1}} \end{aligned}$$



Kapitel 5

Basiskonstruktion

5.1 Motivation für Snapshot-basierte Verfahren

Gemäß Kapitel 4 existieren Konstruktionsverfahren für fehlerfreie Approximationen nur für einfache Probleme. Daher sind empirische Verfahren erforderlich, welche in der Praxis gute Fehlerkonvergenz aufweisen.

Ziel:

- Bestimmung eines „möglichst guten“ RB-Raumes $X_N = \text{span}\{u(\mu^i)\} \subset X$ mit Basis Φ_N welche $\mathcal{M} := \{u(\mu) \mid \mu \in \mathcal{P}\}$ global approximieren.
- Optimales X_N : formalisierbar durch Minimierung eines Funktionals, z.B. minimiere den maximalen Energiefehler

$$\min_{\substack{Y \subset X \\ \dim Y = N}} \max_{\mu \in \mathcal{P}} \| |u(\mu) - u_N(\mu)| \|_{\mu}$$

oder Minimieren des mittleren Projektionsfehlers

$$\min_{\substack{Y \subset X \\ \dim Y = N}} \int_{\mathcal{P}} \| |u(\mu) - P_Y u(\mu)| \|^2 d\mu \quad (5.1)$$

- Gute Basis Φ_N : orthogonal für numerische Stabilität, Hierarchie, so dass Basisvektoren nach Relevanz geordnet sind, d.h. $X_{N'} := \text{span}\{\varphi_1, \dots, \varphi_{N'}\}$ Sequenz von optimalen Räumen für $1 \leq N' \leq N$, damit N' -Variation eine Fehlerkontrolle erlaubt.

Numerische Behandlung:

- „Diskretisierung des Parameterraumes“: Wähle endliche Teilmenge $S_{train} = \{\mu^i\}_{i=1}^n \subset \mathcal{P}$ von Trainingsparametern welche Trainings-Snapshots $M_{train} := \{u(\mu) \mid \mu \in S_{train}\} \subset \mathcal{M}$ definieren.
- Schränke Optimierungsproblem auf S_{train} ein, z.B. statt (5.1)

$$\min_{\substack{Y \subset X \\ \dim Y = N}} \frac{1}{n} \sum_{i=1}^n \|u(\mu^i) - P_Y u(\mu^i)\|^2$$

- Optimierung über Teilräume ist möglicherweise „teuer“, daher eventuell weitere Vereinfachung („Approximation“ der Optimierung, Fehlerschätzer statt Fehler)

Definition 5.1 (Gram-Matrix):

Sei X HR, $\{u_i\}_{i=1}^n \subset X$. Dann ist die Gram-Matrix definiert als $\underline{K} := ((u_i, u_j))_{i,j=1}^n \in \mathbb{R}^{n \times n}$

Lemma 5.2 (Eigenschaften von \underline{K}):

Sei \underline{K} Gram-Matrix von $\{u_i\}_{i=1}^n \subset X$. Dann gilt:

- \underline{K} ist symmetrisch und positiv semidefinit.
- $\text{Rang}(\underline{K}) = \dim \text{span}\{u_i\}_{i=1}^n$
- $\{u_i\}_{i=1}^n$ linear unabhängig $\Leftrightarrow \underline{K}$ positiv definit

Beweis: Übung □

5.2 Gram-Schmidt

Definition 5.3 (Gram-Schmidt Basis):

Sei X HR, $\{u_i\}_{i=1}^n \subset X$ linear unabhängig. Dann ist für $1 \leq m \leq n$ die Gram-Schmidt Basis $\Phi_{GR,m} = \{\varphi_1, \dots, \varphi_m\}$ definiert durch

$$\begin{aligned} \overline{\varphi}_m &:= u_m - \sum_{i=1}^{m-1} (u_m, \varphi_i) \varphi_i \\ \varphi_m &:= \frac{\overline{\varphi}_m}{\|\overline{\varphi}_m\|} \end{aligned}$$

und $X_{GR,m} := \text{span}(\Phi_{GR,m})$ bezeichnet den Gram-Schmidt RB-Raum.

Lemma 5.4 (Eigenschaften von $\Phi_{GR,m}$):

- i) $\Phi_{GR,m}$ ist Orthonormalbasis
- ii) $X_{GR,m} = \text{span}\{u_i\}_{i=1}^m$
- iii) $\max_{j=1,\dots,m} \inf_{v \in X_{GR,m}} \|u_j - v\| = 0$

Beweis:

i)

$$(\varphi_i, \varphi_i) = \left(\frac{\overline{\varphi_i}}{\|\overline{\varphi_i}\|}, \frac{\overline{\varphi_i}}{\|\overline{\varphi_i}\|} \right) = \frac{(\overline{\varphi_i}, \overline{\varphi_i})}{\|\overline{\varphi_i}\|^2} = 1$$

Orthogonalität per Induktion: Induktionsschritt: Sei $(\varphi_i, \varphi_j) = 0$ für alle $j < i$. Dann gilt für $j < i + 1$

$$\begin{aligned} (\overline{\varphi_{i+1}}, \varphi_j) &= (u_{i+1}, \varphi_j) - \sum_{k=1}^{(i+1)-1} (u_{i+1}, \varphi_k) \underbrace{(\varphi_k, \varphi_j)}_{=\delta_{kj}} \\ &= (u_{i+1}, \varphi_j) - (u_{i+1}, \varphi_j) = 0. \end{aligned}$$

Also auch $(\varphi_{i+1}, \varphi_j) = 0$.

- ii) „ \subseteq “ klar nach Konstruktion
- „ \supseteq “ folgt durch Dimensionsbetrachtung

$$\dim \text{span}\{\varphi_i\}_{i=1}^m = m \stackrel{l.u.}{=} \dim \text{span}\{u_i\}_{i=1}^m$$

- iii) nach ii) ist $u_j \in X_{GR,m}, j \leq m$

□

Satz 5.5 (Berechnung von $\Phi_{GR,n}$ über Gram-Matrix):

Sei X HR, $\{u_i\}_{i=1}^n \subset X$ linear unabhängig, $K = ((u_i, u_j))_{i,j=1}^n$ Gram-Matrix mit Cholesky-Zerlegung $K = LL^T$ mit L untere Δ -Matrix mit positiver Diagonalen. Setze $A = (a_{ij})_{i,j=1}^n := (L^T)^{-1}$. Dann ist die Gram-Schmidt-Basis $\Phi_{GR,n} = \{\varphi_1, \dots, \varphi_n\}$ äquivalent berechenbar durch $\varphi_j = \sum_{i=1}^j a_{ij} u_i$ für alle $1 \leq j \leq n$.

Beweis: \Rightarrow Übung.

□

Bemerkung:

- i) Obiger Algorithmus liefert eine orthonormale Basis, garantiert also Stabilität des RB-Verfahrens gemäß Satz 3.6.
- ii) Es existieren nur „triviale“ Approximationsaussagen aus Lemma 5.4 iii). Für Teilbasis $\Phi_{GR,m} = \{\varphi_1, \dots, \varphi_m\}$ werden Snapshots $\{u_i\}_{i=1}^m$ perfekt approximiert, über $\{u_i\}_{i=m+1}^n$ weiß man nichts; schlimmstenfalls sehr schlechte Approximation.
- iii) Die Basis hängt von der Reihenfolge der Snapshots ab, was nur Sinn macht, wenn die $\{u_i\}$ eine natürliche Reihenfolge besitzen.
- iv) GS-Orthonormalisierung häufig als „Post-processing“ für anderweitig erzeugt Basis, z.B. Greedy-Search Basis, Abschnitt 5.4.

5.3 Proper Orthogonal Decomposition (POD)**Satz 5.6 (POD Basis):**

Sei X HR, $\{u_i\}_{i=1}^n \subset X$. Dann ist der Empirische Korrelationsoperator R definiert durch

$$Ru := \frac{1}{n} \sum_{i=1}^n (u_i, u) u_i \quad \text{für } u \in X$$

Es ist $R \in K(X)$ und es existiert orthonormale Menge $\{\varphi_i\}_{i=1}^{n'}$ von $n' < n$ Eigenvektoren zu reellen Eigenwerten $\lambda_1 \geq \lambda_2 \dots \geq \lambda_{n'} > 0$ mit

$$Ru = \sum_{i=1}^{n'} \lambda_i (\varphi_i, u) \varphi_i \quad (5.2)$$

Für $1 \leq m \leq n'$ definieren wir $\Phi_{POD,m} := \{\varphi_i\}_{i=1}^m$ als POD-Basis und $X_{POD,m} = \text{span}(\Phi_{POD,m}) \subset X$ als POD-RB-Raum.

Beweis: R ist linear und beschränkt durch $\|R\| \leq \frac{1}{n} \sum \|u_i\|^2$ und hat endlichdimensionales Bild, also ist R kompakt. R ist selbstadjungiert, denn $(Ru, v) = \frac{1}{n} \sum_{i=1}^n (u_i, u)(u_i, v) = (u, Rv)$ also existiert nach dem Spektralsatz, Satz 2.6, ein entsprechendes endliches Orthonormalsystem, das (5.2) erfüllt. Dieses kann insbesondere nicht unendlich sein, da Bild von R endlichdimensional ist. \square

Bemerkung:

Die Projektion $X \rightarrow X_{POD,m}$ wird in der statistischen Datenanalyse auch *Hotelling-Transformation*, *Principal Component Analysis (PCA)* oder *Karhunen-Loeve Transformation* genannt.

Illustration, Abb. 5.1:

- i) $\{\varphi_i\}_{i=1}^{n'}$ ist orthonormale Basis für $\text{span}\{u_i\}_{i=1}^{n'}$, nicht eindeutig (Spiegelung).
- ii) φ_1 ist Richtung höchster Varianz von $\{u_i\}_{i=1}^n$,
 φ_2 ist Richtung höchster Varianz von $\{P_{X_{POD,1}^\perp} u_i\}$, etc.
- iii) Die Koordinaten der Daten bzgl. der POD-Basis sind unkorreliert.
- iv) $\{\varphi_i\}$ und $\{\sqrt{\lambda_i}\}$ sind die Hauptachsen bzw. Achsenabschnitte des Ellipsoids $(u, R^{-1}u) = 1$.
- v) Falls $X = \mathbb{R}^H$ und $\{u_i\}_{i=1}^n$ normalverteilt mit Mittelwert 0 und Kovarianzmatrix Σ so $R \xrightarrow{P} \Sigma$ für $n \rightarrow \infty$

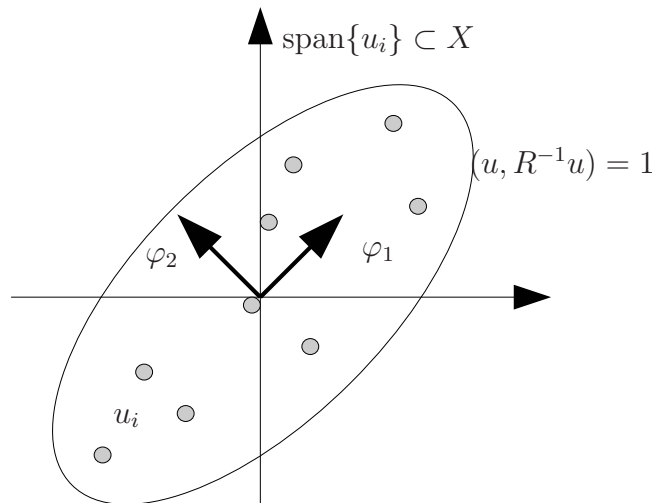


Abbildung 5.1: POD: Ellipsoide aus Kovarianzoperator

Satz 5.7 (Berechnung von $\Phi_{POD,m}$ über Gram-Matrix):

Sei X HR, $\{u_i\}_{i=1}^n \subset X$ und $\underline{K} = ((u_i, u_j))_{i,j=1}^n$ die Gram-Matrix. Dann sind äquivalent:

- i) $\varphi \in X$ ist Eigenvektor von R zu Eigenwert $\lambda > 0$ mit Norm 1 und einer Darstellung $\varphi = \sum a_i u_i$ mit $a \in \text{Kern}(\underline{K})^\perp$.

ii) $a = (a_i)_{i=1}^n \in \mathbb{R}^n$ ist Eigenvektor von $\frac{1}{n}\underline{K}$ zu $\lambda > 0$ mit Norm $\frac{1}{\sqrt{n\lambda}}$.

Beweis: ii) \Rightarrow i)

Sei a EV von $\frac{1}{n}\underline{K}$ zu EW $\lambda > 0$ mit $\|a\| = \frac{1}{\sqrt{n\lambda}}$ also

$$\lambda a = \frac{1}{n}\underline{K}a$$

Multiplikation der i -ten Komponente mit u_i und Summieren ergibt

$$\sum_{i=1}^n u_i \lambda a_i = \sum_{i=1}^n u_i \frac{1}{n} \sum_{j=1}^n (u_i, u_j) a_j$$

Mit $\varphi := \sum u_i a_i$ gilt also

$$\lambda \varphi = \frac{1}{n} \sum_{i=1}^n u_i (u_i, \varphi) = R\varphi$$

also φ EV von R zu EW λ . Für die Norm folgt

$$\begin{aligned} \|\varphi\|^2 &= \left(\sum_{i=1}^n a_i u_i, \sum_{j=1}^n a_j u_j \right) = \sum a_i a_j (u_i, u_j) = a^T \underline{K} a = a^T n \lambda a \\ &= n \lambda \|a\|^2 = n \lambda \frac{1}{\sqrt{n\lambda}^2} = 1 \end{aligned}$$

\underline{K} ist symmetrisch, also existiert vollständiges Orthonormalsystem von Eigenvektoren. $\text{Kern}(\underline{K})$ wird aufgespannt von Eigenvektoren zu Eigenwert 0, also $a \perp \text{Kern}(\underline{K})$, da a EV zu $\lambda > 0$.

i) \Rightarrow ii):

Sei φ Eigenvektor von R zu EW $\lambda > 0$ und $\|\varphi\| = 1$. Sei $\bar{a} \in \mathbb{R}^n$ mit $\varphi = \sum \bar{a}_i u_i$ (existiert, weil $\varphi \in \text{Bild}(R)$). Verschiebungen von \bar{a} um Vektoren $a^o \in \text{Kern}(\underline{K})$ erhalten φ :

$$\begin{aligned} \varphi' := \sum_{i=1}^n (\bar{a}_i + a_i^o) u_i \Rightarrow (\varphi', u_k) &= \left(\sum_i \bar{a}_i u_i, u_k \right) + \left(\sum_i a_i^o u_i, u_k \right) \\ &= (\varphi, u_k) + \underbrace{\sum_i a_i^o (u_i, u_k)}_{=\underline{K} \cdot a^o = 0} \end{aligned}$$

Wähle speziell $a := \bar{a} - P\bar{a}$, P orthogonale Projektion auf $\text{Kern}(\underline{K})$, also $a \in \text{Kern}(\underline{K})^\perp$, $P\bar{a} \in \text{Kern}(\underline{K}) \Rightarrow \varphi = \sum_{i=1}^n \bar{a}_i u_i = \sum_{i=1}^n a_i u_i$. Da φ EV zu

$\lambda > 0$ gilt

$$\frac{1}{n} \sum_{i=1}^n (u_i, \sum_{j=1}^n a_j u_j) u_i = \lambda \sum_{j=1}^n a_j u_j$$

Testen mit u_k liefert

$$\frac{1}{n} \underbrace{\sum_{i,j=1}^n (u_i, u_j) a_j (u_i, u_k)}_{(\underline{K}^2 a)_k} = \lambda \underbrace{\sum_{j=1}^n a_j (u_j, u_k)}_{(\underline{K} a)_k}$$

Also $\frac{1}{n} \underline{K}^2 a = \underline{K} a$, also $\underline{K} a$ EV von $\frac{1}{n} \underline{K}$ zu EW λ . Dann ist schon a EV, da $a \in \text{Kern}(\underline{K})^\perp$:

Sei $a = \sum_{i=1}^{n'} \mu_i a^i$ mit a^i EV von $\frac{1}{n} \underline{K}$ zu $\lambda_i > 0$

$$\Rightarrow \underline{K} a = n \cdot \sum_{i=1}^{n'} \mu_i \lambda_i a^i \Rightarrow \frac{1}{n} \underline{K}^2 a = n \sum_{i=1}^{n'} \mu_i \lambda_i^2 a^i$$

Damit $\underline{K} a$ EV muss also $\mu_i = 0$ für $\lambda_i \neq \lambda$ sein

$$\Rightarrow a = \sum_{i:\lambda_i=\lambda} \mu_i a^i \Rightarrow \frac{1}{n} \underline{K} a = \sum_{i:\lambda_i=\lambda} \mu_i \lambda_i a^i = \lambda \cdot a$$

Wie im ersten Teil gilt für die Norm:

$$1 = \|\varphi\|^2 = a^T \underline{K} a = n \lambda a^T a = n \lambda \|a\|^2 \quad \text{also} \quad \|a\| = \frac{1}{\sqrt{n \lambda}}$$

□

Bemerkung:

Falls X endlichdimensional mit $\dim X = H$, kann daher die POD entweder als teures Eigenwertproblem für R in X (Komplexität $O(H^3)$) oder, meist günstiger, als Eigenwertproblem für \underline{K} (Komplexität $O(n^3)$) ermittelt werden. “Method of Snapshots” oder “Kernel PCA”.

Die POD kann manchmal auch über eine Singulärwertzerlegung der Koeffizientenmatrix berechnet werden:

Lemma 5.8 (Berechnung für $X = \mathbb{R}^H$ über SVD):

Sei $X = \mathbb{R}^H$, $U = [u_1, \dots, u_n] \in \mathbb{R}^{H \times n}$ Snapshotmatrix mit Rang n' und $U = \Phi S V^T$ eine (verkürzte) Singulärwertzerlegung mit $\Phi \in \mathbb{R}^{H \times n'}$, orthonormale Spalten, $S = \text{diag}(\sigma_1, \dots, \sigma_{n'}) \in \mathbb{R}^{n' \times n'}$ und $V \in \mathbb{R}^{n \times n'}$, orthonormale Spalten. Falls $\sigma_1 > \sigma_2 > \dots > 0$ echt fallend, so ist $\Phi_{POD, n'} = \Phi$.

Beweis: Sei $\Phi = (\overline{\varphi}_1, \dots, \overline{\varphi}_{n'})$. Nach Definition gilt $Ru = \frac{1}{n}UU^T u, \forall u \in X$. Damit ist $\overline{\varphi}_i$ Eigenvektor von R :

$$R\overline{\varphi}_i = \frac{1}{n}UU^T\overline{\varphi}_i = \frac{1}{n}\Phi S \underbrace{V^T V}_{=I_{n'}} S^T \underbrace{\Phi^T \overline{\varphi}_i}_{=\mathbf{e}_i} = \frac{1}{n}\sigma_i^2 \overline{\varphi}_i.$$

Die Eigenwerte $\frac{1}{n}\sigma_i^2$ sind monoton fallend, also identisch sortiert wie Spektralzerlegung von R , d.h. $\frac{1}{n}\sigma_i^2 = \lambda_i$ und $\overline{\varphi}_i = \varphi_i$ oder $\overline{\varphi}_i = -\varphi_i$. \square

Satz 5.9 (Approximationsfehler für $X_{POD,m}$):

Sei X HR, $\{u_i\}_{i=1}^n \subset X$ Teilmenge und für $Y \subset X$ Teilraum sei der mittlere quadratische Projektionsfehler definiert durch

$$J(Y) := \frac{1}{n} \sum_{i=1}^n \|u_i - P_Y u_i\|^2.$$

Dann gilt für den POD-RB-Raum $X_{POD,m}$ gemäß Satz 5.6

$$J(X_{POD,m}) = \sum_{i=m+1}^{n'} \lambda_i$$

mit $\{\lambda_i\}_{i=1}^{n'}$ Eigenwerten von R .

Beweis: Sei $\psi = \{\psi_1, \dots, \psi_m\}$ orthonormale Basis von Y . Dann folgt

$$\begin{aligned} J(Y) &= \frac{1}{n} \sum_{i=1}^n \|u_i - P_Y u_i\|^2 = \frac{1}{n} \sum_{i=1}^n \|u_i - \sum_{j=1}^m (u_i, \psi_j) \psi_j\|^2 \\ &= \frac{1}{n} \sum_{i=1}^n \|u_i\|^2 - \frac{2}{n} \sum_{i=1}^n \sum_{j=1}^m (u_i, \psi_j)^2 + \frac{1}{n} \sum_{i=1}^n \sum_{j,k=1}^m (u_i, \psi_j)(u_i, \psi_k) \underbrace{(\psi_j, \psi_k)}_{=\delta_{jk}} \\ &= \frac{1}{n} \sum_{i=1}^n \|u_i\|^2 - \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^m (u_i, \psi_j)^2. \end{aligned}$$

Wegen $u_i \in \text{Bild}(R) = X_{POD,n'}$ ist

$$u_i = \sum_{j=1}^{n'} \varphi_j(\varphi_j, u_i), \quad \|u_i\|^2 = \sum_{j=1}^{n'} (\varphi_j, u_i)^2$$

also der mittlere Projektionsfehler

$$\begin{aligned}
J(X_{POD,m}) &= \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^{n'} (\varphi_j, u_i)^2 - \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^m (u_i, \varphi_j)^2 \\
&= \frac{1}{n} \sum_{i=1}^n \sum_{j=m+1}^{n'} (\varphi_j, u_i)^2 \\
&= \sum_{j=m+1}^{n'} \left(\frac{1}{n} \sum_{i=1}^n (\varphi_j, u_i) u_i, \varphi_j \right) \\
&= \sum_{j=m+1}^{n'} (R\varphi_j, \varphi_j) = \sum_{j=m+1}^{n'} \lambda_j(\varphi_j, \varphi_j) = \sum_{j=m+1}^{n'} \lambda_j.
\end{aligned}$$

□

Satz 5.10 (Bestapproximation durch $X_{POD,m}$):

Unter allen Räumen der Dimension m ist $X_{POD,m}$ bzgl. J optimal.

$$J(X_{POD,m}) = \inf_{\substack{Y \subset X \\ \dim Y = m}} J(Y)$$

Beweis: Übung

□

Bemerkung:

- i) POD liefert also orthonormierte Basis, garantiert Stabilität des RB-Verfahrens, Satz 3.6.
- ii) Es existieren Approximationsaussagen bzgl. mittleren quadratischen Fehler, Satz 5.9, sogar Optimalität nachweisbar. Die POD Teil-Basen ermöglichen eine Approximation aller Snapshots mit Fehlerkontrolle.
- iii) Die Basis hängt nicht von der Reihenfolge der Snapshots ab.
- iv) Die POD-Basisgenerierung kann auch zur Erweiterung einer gegebenen ONB Φ verwendet werden, indem $\{u_i\}_{i=1}^n$ zunächst bezüglich Φ orthogonalisiert werden zu $\{\tilde{u}_i\}_{i=1}^n$ mit $\tilde{u}_i := u_i - P_{\text{span}(\Phi)}u_i$ und eine Basis $\tilde{\Phi}_{POD,m}$ hierfür berechnet wird. Dann ist $\Phi \cup \tilde{\Phi}_{POD,m}$ neue ONB.
- v) Die POD-Teilbasen sind Hierarchische Basen, d.h. $\Phi_m \subset \Phi_{m+1}$.

5.4 Greedy Search

Definition 5.11 (Basis durch „Greedy Search“):

Sei X HR, $S_{train} = \{\mu^i\}_{i=1}^n \subset \mathcal{P}$ Menge von Trainingsparametern. Sei $\Delta(Y, \mu) \in \mathbb{R}^+$ für Teilräume $Y \subset X$ und Parameter $\mu \in \mathcal{P}$ und $\varepsilon_{tol} > 0$ eine Fehler-toleranz. Die Greedy-Basis $\Phi_{GRE,m}$, Greedy-RB-Raum $X_{GRE,m}$ und Sample-Menge S_m für $m = 1, \dots, N$ sind iterativ definiert durch

$$X_{GRE,0} := \{0\}, S_0 := \emptyset, m = 0, \Phi_{GRE,0} := \emptyset.$$

$$\text{Solange } \varepsilon_m := \max_{\mu \in S_{train}} \Delta(X_{GRE,m}, \mu) > \varepsilon_{tol}$$

$$m \mapsto m + 1$$

$$\mu_m^* := \arg \max \Delta(X_{GRE,m-1}, \mu)$$

$$S_m := S_{m-1} \cup \{\mu_m^*\}$$

$$\varphi_m := u(\mu_m^*), \Phi_{GRE,m} := \Phi_{GRE,m-1} \cup \{\varphi_m\}$$

$$X_{GRE,m} := X_{GRE,m-1} + \text{span}(\varphi_m)$$

$$\text{Setze zuletzt } N := m.$$

Lemma 5.12 (Fehlerindikatoren, Terminieren der Suche):

i) Falls für alle $\mu \in \mathcal{P}, Y \subset X$ gilt

$$u(\mu) \in Y \Rightarrow \Delta(Y, \mu) = 0$$

so terminiert obiger Algorithmus mit $N \leq n$ und

$$\max_{\mu \in S_{train}} \Delta(X_{GRE,N}, \mu) \leq \varepsilon_{tol}$$

ii) dies erfüllen z.B.

$$\Delta(Y, \mu) := \|u(\mu) - u_N(\mu)\|$$

$$\Delta(Y, \mu) := \|u(\mu) - P_Y u(\mu)\|$$

$$\Delta(Y, \mu) := \Delta_N^{en}(\mu)$$

oder andere Fehlerschätzer aus Satz 3.13, wobei $X_N = Y$ gesetzt wird.

Beweis: Siehe Übung. □

Lemma 5.13 (Fehlerraussage):

Für $\|u(\mu) - u_{N'}(\mu)\|$ oder Δ_N als Fehlerindikator gilt für $1 \leq N' \leq N$ und $X_{N'} := X_{GRE, N'}$

$$\max_{\mu \in S_{train}} \|u(\mu) - u_{N'}(\mu)\| \leq \varepsilon_{N'}.$$

Beweis: Klar nach Konstruktion, da $\|u(\mu) - u_{N'}(\mu)\| \leq \Delta(X_{GRE, N'}, \mu)$ \square

Bemerkung:

- i) $\Phi_{GRE, m}$ ist eine Lagrange-RB-Basis zur Sample-Menge S_m , i.A. nicht orthonormal, kann für numerische Stabilität mit Gram-Schmidt orthonormalisiert werden.
- ii) Die „Greedy Search“ ist also ein akkumulatives Basisgenerierungsverfahren, welches iterativ den „schlechtest“-aufgelösten Parameter μ_m^* wählt und $u(\mu_m^*)$ berechnet und als neuen Basisvektor wählt. insofern kann dies als approximative Lösung des Optimierungsproblems

$$\min_{\substack{Y \subset X \\ \dim Y = N}} \max_{\mu \in \mathcal{P}} \Delta(Y, \mu)$$

interpretiert werden: statt maximieren über \mathcal{P} maximieren über S_{train} , statt minimieren über $Y \subset X, \dim Y = N \rightsquigarrow$ iterative Sequenz von Räumen $Y = X_{GRE, m}$.

- iii) „Greedy“ Algorithmus ist hervorragendes Einsatzfeld für Fehlerschätzer aus Satz 3.13: $\Delta(Y, \mu)$ kann sehr schnell für alle $\mu \in S_{train}$ ausgerechnet werden ohne dass Berechnung aller $u(\mu)$ erfolgen muss (im Gegensatz zu POD, Gram-Schmidt oder Greedy mit X -Norm oder Projektionsfehler als Fehlerindikator). Dadurch können sehr große Mengen $S_{train} \subset \mathcal{P}$ gewählt werden. Dies erhöht die Erwartung, dass $\Phi_{GRE, N}$ auch für neue Parameter $\mu \in \mathcal{P} \setminus S_{train}$ eine gute RB-Approximation liefert.
- iv) Man kann eine Greedy-Search auch zur Erweiterung einer vorhandenen Basis $\Phi \subset X$ verwenden.
- v) Ist S_{train} zu klein, kann das RB-Modell Overfitting aufweisen, d.h.

$$\sup_{\mu \in \mathcal{P}} \|u(\mu) - u_N(\mu)\| \gg \varepsilon_{tol}.$$

Man kann als Verbesserung das Verfahren zu einer adaptiven Methode erweitern, indem die S_{train} -Menge schrittweise verfeinert wird in Parameterbereichen, die schlecht aufgelöst sind. (Adaptive Parameter-Gitter, „Multi-Stage Greedy“)

- vi) Anwendung des „Greedy“ Algorithmus mit e_N^s oder $\Delta_N^s, \Delta_N^{s,rel}$ als Fehlerindikator kann eine sehr kleine Basis erzeugen, welche die Ausgabe gut approximiert. Die Lösung $u(\mu)$ ist dann nicht unbedingt sehr gut approximiert. Bei Verwendung von e_N oder $\Delta_N, \Delta_N^{rel}, \Delta_N^{en}, \Delta_N^{en,rel}$ als Fehlerindikator wird i.a. eine große Basis erzeugt welche dann sowohl die Feldvariable $u(\mu)$ als auch die Ausgabe gut approximiert.
- vii) Die Greedy-Basis hängt meistens nicht von der Reihenfolge der Snapshots ab, nur falls zufällig der maximale Fehlerindikator für mehrere Snapshots realisiert wird gibt es eine Uneindeutigkeit des $\arg \max$, bei der „Wahl“ eines maximierenden Elementes kann also die Reihenfolge eine Rolle spielen („wähle erstes Snapshot, das maximalen Fehler erzeugt“).
- viii) Die Greedy-Suche erzeugt eine hierarchische Basis, welche nichttriviale Fehleraussagen erlaubt: Für $1 \leq N' \leq N$ gilt nach Konstruktion unter Verwendung von $\Delta(Y, \mu) := \|u(\mu) - u_{N'}(\mu)\|$ oder $\Delta(Y, \mu) := \Delta_{N'}(\mu)$

$$\max_{\mu \in S_{train}} \|u(\mu) - u_{N'}(\mu)\| \leq \varepsilon_{N'}$$

Satz 5.14 (Konvergenzraten für Greedy-Verfahren [BCD⁺10]):

Sei $\mathcal{M} := \{u(\mu) | \mu \in \mathcal{P}\}$, \mathcal{P} kompakt, und es gelten die Voraussetzungen von (P_N) . Sei $\sigma_N := \max_{\mu \in \mathcal{P}} \Delta_N(\mu)$ für $N \in \mathbb{N}$.

- i) (Algebraische Konvergenz) Falls $d_n(\mathcal{M}) \leq Mn^{-\alpha}$ für gewisse $\alpha, M > 0$ und alle $n \in \mathbb{N}$, und $d_0(\mathcal{M}) \leq M$ dann gilt

$$\sigma_n \leq CMn^{-\alpha}, n > 0$$

mit geeigneter (explizit berechenbarer) Konstante $C > 0$.

- ii) (Exponentielle Konvergenz) Falls $d_n(\mathcal{M}) \leq Me^{-an^\alpha}$ für $n \geq 0, M, a, \alpha > 0$, dann gilt

$$\sigma_n \leq CM e^{-cn^\beta}, n \geq 0$$

für $\beta := \alpha/(\alpha + 1)$ und geeigneten (explizit berechenbaren) Konstanten $c, C > 0$.

Bemerkung:

- Das Greedy-Verfahren ist also höchstens um einige Konstanten schlechter als die Bestapproximation mit linearen Unterräumen.
- Das obige ist also eine Art “idealisiertes” Greedy-Verfahren (Suche über \mathcal{P} statt endlicher Teilmenge). In derselben Arbeit wurde auch ein ähnliches Ergebnis (“Robustheit”) angegeben, welches algorithmische Approximationen berücksichtigt.

Kapitel 6

Allgemeine Parameterabhängigkeit

6.1 Empirische Interpolation (EI)

Motivation:

- Kapitel 3 zeigte Wichtigkeit von „separabler Parameterabhängigkeit“ für effiziente Offline/Online Zerlegung und Glattheit der Lösung bzgl. μ .
- Gesucht ist daher ein Approximationsverfahren für parametrische Funktionen $g : \Omega \times \mathcal{P} \rightarrow \mathbb{R}$ der Form

$$g(x; \mu) \approx I_M(g(\cdot; \mu))(x) = \sum_{m=1}^M \Theta_g^m(\mu) g^m(x)$$

mit skalaren Funktionen $\Theta_g^m(\mu)$ und einer “kollateralen Reduzierten Basis” $Q_M := \{g^m\}_{m=1}^M$

- Statt allgemein approximierende Räume (zu hohe Dimension) oder Taylor-Ansatz (nur lokale Approximation) wird wieder „Snapshot“-basierter Ansatz gewählt, d.h. $Q_M \subset \text{span}\{g(\cdot; \mu) \mid \mu \in S_{train} \subset \mathcal{P}\}$.
- Die „Empirische Interpolation“ ist eine Möglichkeit. Details zu dem Verfahren finden sich in [BMNP04, MNPP07]
 - Y. Maday, N.C. Nguyen, A.T. Patera, G.S.H. Pau, *A general, multi-purpose interpolation procedure: the magic points*, Preprint RO7037, Laboratoire Jacques-Louis-Lions, Université Pierre et Marie Curie, Paris, 2007

Definition 6.1 (Empirische Interpolation):

Sei $G \subset C^0(\bar{\Omega}, \mathbb{R})$ Menge von zu interpolierenden Funktionen. Für $M \in \mathbb{N}$, $M \leq \dim(\text{span}(G))$ definieren wir rekursiv die Interpolationspunktmenge $T_M \subset \bar{\Omega}$ und Kollaterale Reduzierte Basis $Q_M \subset \text{span}(G)$:

$$\begin{aligned}
 M = 1 : \quad \tilde{q}_1 &:= \arg \max_{g \in G} \|g\|_\infty \\
 x_1 &:= \arg \max_{x \in \bar{\Omega}} |\tilde{q}_1(x)|, \quad T_1 := \{x_1\} \\
 q_1 &:= \tilde{q}_1 / \tilde{q}_1(x_1), \quad Q_1 := \{q_1\} \\
 M > 1 : \quad \tilde{q}_M &:= \arg \max_{g \in G} \|g - I_{M-1}(g)\|_\infty \\
 r_M &:= \tilde{q}_M - I_{M-1}(\tilde{q}_M) \\
 x_M &:= \arg \max_{x \in \bar{\Omega}} |r_M(x)|, \quad T_M := T_{M-1} \cup \{x_M\} \\
 q_M &:= r_M / r_M(x_M), \quad Q_M := Q_{M-1} \cup \{q_M\}
 \end{aligned}$$

wobei $I_M : C^0(\bar{\Omega}, \mathbb{R}) \rightarrow \text{span}(Q_M)$ den Interpolationsoperator zu den Interpolationspunkten T_M bezeichnet, d.h. $I_M(g)(x_i) = g(x_i) \quad \forall g \in C^0(\bar{\Omega}, \mathbb{R}), i = 1, \dots, M$.

Bemerkung:

- i) In der Praxis werden die obigen Optimierungsprobleme zur Bestimmung von \tilde{q}_m und x_m durch eine einfache lineare Suche realisiert, indem endliches $\bar{\Omega}$ und G gewählt werden.
- ii) Es sind Mehrdeutigkeiten von \tilde{q}_m und x_m möglich, welche im Fall $\bar{\Omega}, G$ endlich durch eine Aufzählung der Mengen und „Wahl des Maximierers mit kleinstem Index“ eindeutig werden.
- iii) Die Basis Q_M ist weder orthogonal noch nodal, aber hierarchisch, d.h. $Q_{M-1} \subset Q_M$ und die Basisfunktionen sind beschränkt durch

$$1 = q_m(x_m) \geq \sup_{x \in \bar{\Omega}} |q_m(x)|.$$

Für analytische Untersuchungen wird später die nodale Basis $\xi_M \subset \text{span}(Q_M)$ zu den Knoten T_M betrachtet, welche jedoch nicht mehr hierarchisch ist, d.h. $\xi_{M-1} \not\subset \xi_M$.

- iv) Die Erzeugung von Q_M und T_M kann als „Greedy“ Minimierungsalgorithmus zur Minimierung von

$$\min_{\substack{X_M \subset \text{span}(G) \\ \dim X_M = M \\ T_M \subset \bar{\Omega} \\ |T_M| = M}} \max_{g \in G} \|g - I_M(g)\|_\infty$$

interpretiert werden. Es kann jedoch kein monotoner Fehlerabfall garantiert werden.

Lemma 6.2 (Berechnung der Interpolation):

Seien $T_M = \{x_1, \dots, x_M\}$, $Q_M = \{q_1, \dots, q_M\}$ gemäß Konstruktion aus Definition 6.1 gegeben. Dann ist die Matrix $\underline{Q}_M := (q_j(x_i))_{i,j=1}^M \in \mathbb{R}^{M \times M}$ untere Dreiecksmatrix mit 1 auf der Diagonalen, also regulär. Sei $g \in C^0(\bar{\Omega}, \mathbb{R})$, $\underline{g}_M := (g(x_i))_{i=1}^M \in \mathbb{R}^M$ und $\underline{\alpha}_M = (\alpha_i)_{i=1}^M \in \mathbb{R}^M$ Lösung von

$$\underline{Q}_M \underline{\alpha}_M = \underline{g}_M$$

Dann ist die Interpolierte von g gerade

$$I_M(g) = \sum_{i=1}^M \alpha_i q_i \quad (6.1)$$

Beweis: Seien $i, j = 1, \dots, M$:

$$\begin{aligned} i = j : \quad (\underline{Q}_M)_{ij} &= q_i(x_i) = r_i(x_i)/r_i(x_i) = 1 \\ j > i : \quad (\underline{Q}_M)_{ij} &= q_j(x_i) = \frac{1}{r_j(x_j)} r_j(x_i) = 0, \end{aligned}$$

weil $r_j(x_i) = \tilde{q}_j(x_i) - I_{j-1}(\tilde{q}_j)(x_i) = 0$, da I_{j-1} Interpolierende zu T_{j-1} und $x_i \in T_{j-1}$. Also \underline{Q}_M untere Δ -Matrix, 1 auf der Diagonalen. Für (6.1) reicht es zu zeigen, dass beide Seiten in den Punkten T_M übereinstimmen:

$$\sum_{j=1}^M \alpha_j q_j(x_i) = \sum_{j=1}^M (\underline{Q}_M)_{ij} \alpha_j = (\underline{g}_M)_i = g(x_i) = I_M(g)(x_i).$$

□

Beispiel 6.3 (EI für Polynome):

Sei $G = \{1, x, x^2\}$ Monome auf $\bar{\Omega} = [-1, 1]$, siehe Abb. 6.1. Dann ist

$$\tilde{q}_1 := \arg \max_{g \in G} \|g\|_\infty \text{ beliebig, z.B. } \tilde{q}_1(x) = 1,$$

und

$$x_1 := \arg \max_{x \in \bar{\Omega}} |\tilde{q}_1(x)| \text{ beliebig, z.B. } x_1 = -1.$$

Somit ergibt sich $q_1(x) = \tilde{q}_1(x)/\tilde{q}_1(x_1) = 1$.

Dann ist

$$\tilde{q}_2 := \arg \max_{g \in G} \|g - I_1(g)\|_\infty = x, \quad r_2 = \tilde{q}_2 - I_1(\tilde{q}_2) = x + 1,$$

$$x_2 = \arg \max_{x \in \bar{\Omega}} |r_2(x)| = 1$$

und es ergibt sich $q_2(x) = r_2(x)/r_2(x_2) = \frac{1}{2}(x + 1)$.

Schließlich ist

$$\tilde{q}_3 := \arg \max_{g \in G} \|g - I_2(g)\|_\infty = x^2, \quad r_3(x) = x^2 - I_2(\tilde{q}_3) = x^2 - 1,$$

$$x_3 = \arg \max_{x \in \bar{\Omega}} |r_3(x)| = 0$$

und es ergibt sich $q_3(x) = r_3(x)/r_3(x_3) = 1 - x^2$.

i	$\ x^i\ _\infty$	$\ x^i - I_1(x^i)\ _\infty$	$\ x^i - I_2(x^i)\ _\infty$
0	1	0	0
1	1	2	0
2	1	1	1

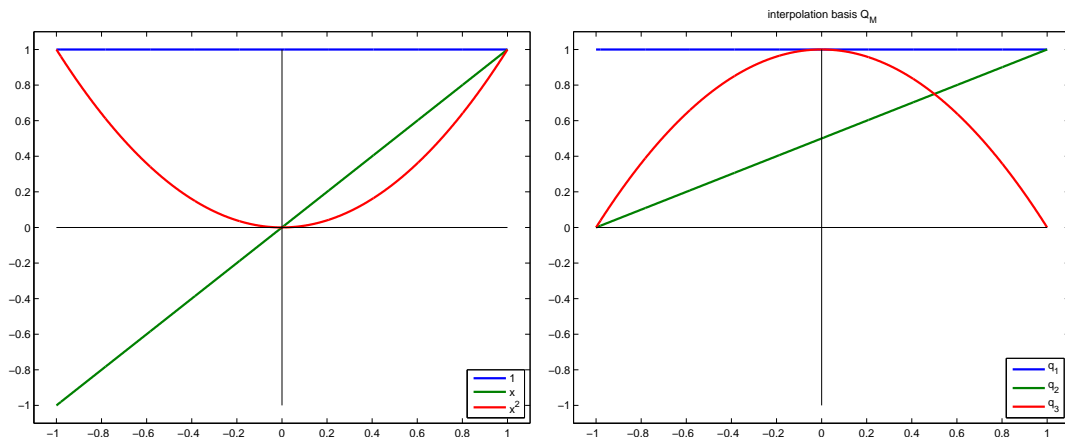


Abbildung 6.1: EI von 3 Polynomen.

Größeres Beispiel: $G = \{x^i\}_{i=0}^{29}$, siehe Abb. 6.2. Erkenntnisse:

- alle q_j beschränkt $\|q_i\|_\infty = 1$
- $q_j(x_i) = 0$ $i < j$ wegen Interpolation
- $q_j(x_j) = 1$

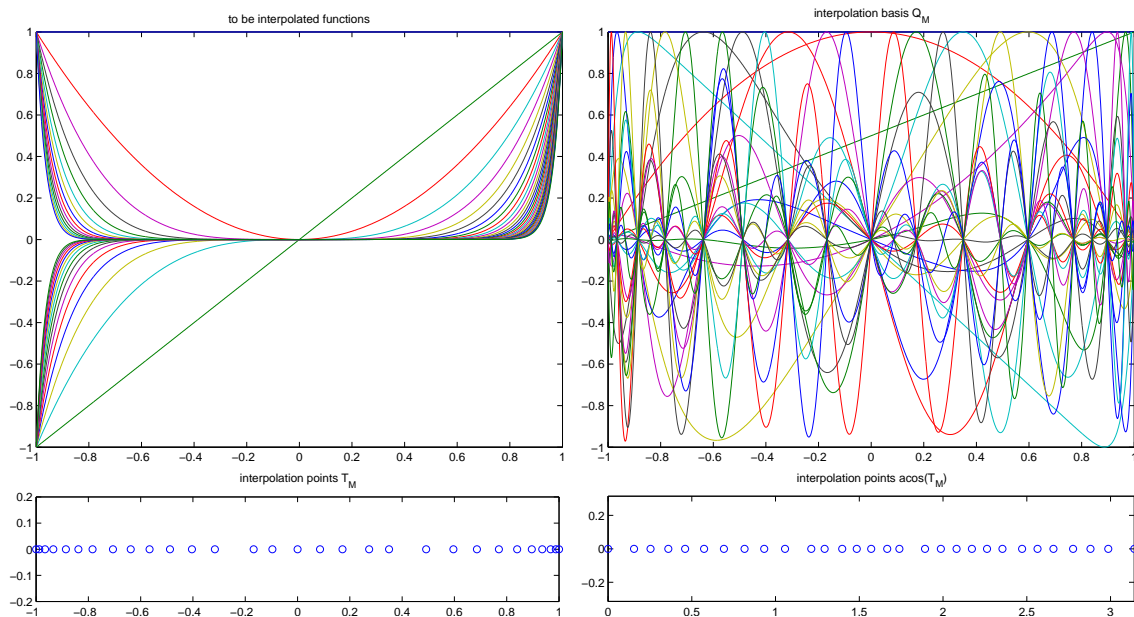


Abbildung 6.2: Empirische Interpolation für Polynome

- $\cos^{-1}(T_M)$ ist etwa „äquidistant“
- T_M approximieren die Chebyshev-Knoten, die für polynomiale Interpolation als beste Wahl bekannt sind!

$\Rightarrow T_M$ sind sogenannte „magic points“.

6.2 Eigenschaften der EI

Satz 6.4 (Lebesgue-Konstante):

Sei $I_M : C^0(\bar{\Omega}, \mathbb{R}) \rightarrow X_M = \text{span}\{\xi_i\}_{i=1}^M \subset C^0(\bar{\Omega}, \mathbb{R})$ Interpolationsoperator zu den Punkten $\{x_i\}_{i=1}^M \subset \bar{\Omega}$ und $\{\xi_i\}_{i=1}^M$ eine nodale Basis, d.h. $\xi_i(x_j) = \delta_{ij}$, $I_M = \sum_{i=1}^M u(x_i)\xi_i$. Dann ist $\Lambda_M := \max_{x \in \bar{\Omega}} \sum_{i=1}^M |\xi_i(x)|$ die Lebesgue-Konstante der Interpolation.

i) Es gilt

$$\|u - I_M(u)\|_\infty \leq (1 + \Lambda_M) \inf_{v \in X_M} \|u - v\|_\infty \quad \forall u \in C^0(\bar{\Omega}, \mathbb{R}).$$

ii) Für die Empirische Interpolation gilt

$$\Lambda_M \leq 2^M - 1.$$

Beweis:

i) Sei $u \in C^0(\bar{\Omega}, \mathbb{R})$, $x \in \bar{\Omega}$ und $v = \sum a_i \xi_i \in X_M$. Dann ist

$$\begin{aligned} |u(x) - I_M(u)(x)| &= \left| u(x) - \underbrace{\sum_{i=1}^M a_i \xi_i(x)}_{=v(x)} + \sum_{i=1}^M a_i \xi_i(x) - \underbrace{\sum_{i=1}^M u(x_i) \xi_i(x)}_{=I_M(u)} \right| \\ &\leq |u(x) - v(x)| + \left| \sum_{i=1}^M \xi_i(x) (a_i - u(x_i)) \right| \end{aligned} \quad (6.2)$$

Für den letzten Term gilt wegen $\{\xi_i\}$ nodal:

$$\begin{aligned} \left| \sum_{i=1}^M \xi_i(x) (a_i - u(x_i)) \right| &= \left| \sum_{i=1}^M \xi_i(x) \left(\sum_{j=1}^M a_j \xi_j(x_i) - u(x_i) \right) \right| \\ &\leq \sum_{i=1}^M |\xi_i(x)| |v(x_i) - u(x_i)| \\ &\leq \sum_{i=1}^M |\xi_i(x)| \|u - v\|_\infty \\ &\leq \max_{x \in \bar{\Omega}} \sum_{i=1}^M |\xi_i(x)| \|u - v\|_\infty. \end{aligned}$$

Also für (6.2):

$$|u(x) - I_M(u)(x)| \leq \|u - v\|_\infty + \Lambda_M \|u - v\|_\infty$$

Nach Infimum über $v \in X_M$ und Supremum über $x \in \bar{\Omega}$ folgt i).

ii) \Rightarrow Übung.

□

Bemerkung:

Obige Abschätzung ist sehr pessimistisch, in der Praxis werden meist bessere Lebesgue-Konstanten beobachtet, jedoch ist die Schranke scharf, d.h. es existieren Beispiele mit $\Lambda_M = 2^M - 1$.

Satz 6.5 (Exponentielle Konvergenz des Interpolationsfehlers):

Falls es eine Sequenz von exponentiell approximierenden Unterräumen gibt,

d.h. $Z_1 \subset Z_2 \subset \dots \subset Z_M \subset \dots \subset \text{span}(G)$, $\dim Z_M = M$ und existiert ein $c > 0, \alpha > \log(4)$ s.d.

$$\inf_{v \in Z_M} \|u - v\|_\infty \leq c \cdot e^{-\alpha M} \quad \forall u \in G, M \in \mathbb{N}$$

Dann findet der EI-Basiskonstruktionsprozess aus Definition 6.1 „fast so gute“ Räume, indem

$$\|u - I_M(u)\|_\infty \leq c \cdot e^{-(\alpha - \log 4)M}$$

Beweis: Siehe Maday et al. [MNPP07]. □

Bemerkung:

Anschaulich ist die obige Forderung an die Daten eine „Beschränktheit durch eine Müsli-Packung“, einem Polytop mit exponentiell abnehmenden Seitenlängen.

Satz 6.6 (A-posteriori Fehlerschätzer für EI):

Seien $I_M, I_{M'} : C^0(\bar{\Omega}, \mathbb{R}) \rightarrow \text{span}(G)$ EI-Operatoren für $M' > M$, $Q_M \subset Q_{M'} = \{q_i\}_{i=1}^{M'}$, $T_M \subset T_{M'} = \{x_i\}_{i=1}^{M'}$. Sei $\underline{Q} := (q_j(x_i))_{i,j=M+1}^{M'}$. Für $g \in G$ sei $\underline{g}' := (g(x_i) - I_M(g)(x_i))_{i=M+1}^{M'}$ und $\underline{\alpha}' = \{\alpha'_i\}_{i=M+1}^{M'} = \underline{Q}^{-1} \underline{g}'$. Falls $g \in \text{span}(Q_{M'})$, so gelten die folgenden a-posteriori Schranken für den Interpolationsfehler:

$$\begin{aligned} \|g - I_M(g)\|_\infty &\leq \Delta_{M,M',\infty}(g) := \|\underline{\alpha}'\|_1 = \sum_{i=M+1}^{M'} |\alpha'_i|, \\ \|g - I_M(g)\|_{L^2} &\leq \Delta_{M,M',2}(g) := ((\underline{\alpha}')^T \underline{K}_Q \underline{\alpha}')^{1/2} \end{aligned}$$

mit $\underline{K}_Q := (\int_\Omega q_i q_j)_{i,j=M+1}^{M'}$.

Beweis: Nach Definition der EI gilt

$$I_M(g) = \sum_{i=1}^M \alpha_i q_i, \quad I_{M'}(g) = \sum_{i=1}^{M'} \alpha'_i q_i,$$

mit $\underline{Q}_M \underline{\alpha}_M = \underline{g}_M$ und $\underline{Q}_{M'} \underline{\alpha}_{M'} = \underline{g}_{M'}$. Da \underline{Q}_M oberer linker Block, \underline{Q} unterer rechter Block von $\underline{Q}_{M'}$ und jeweils untere Dreiecksmatrizen sind gilt $\alpha_i = \alpha'_i$ für $i = 1, \dots, M$ und

$$g - I_M(g) = I_{M'}(g) - I_M(g) = \sum_{i=1}^{M'} \alpha'_i q_i - \sum_{i=1}^M \alpha_i q_i = \sum_{i=M+1}^{M'} \alpha'_i q_i.$$

Für die L^∞ Norm folgt

$$\|g - I_M(g)\|_\infty \leq \sum_{i=M+1}^{M'} |\alpha'_i| \|q_i\|_\infty \leq \sum_{i=M+1}^{M'} |\alpha'_i| = \Delta_{M,M',\infty}(g).$$

da $\|q_i\|_\infty = 1$. Für die L^2 -Norm des Interpolationsfehlers gilt

$$\|g - I_M(g)\|_{L^2} = \left(\sum_{i,j=M+1}^{M'} \alpha'_i \alpha'_j \int_\Omega q_i q_j \right)^{1/2} = \Delta_{M,M',2}(g).$$

□

6.3 Anwendung für lineare RB-Methoden

Falls eine Linearform in (P) keine separable Parameterabhängigkeit aufweist, liefert die EI eine Approximation:

Definition 6.7 (EI für Funktionale):

Sei $f(v; \mu) = \int_\Omega g(x, \mu) v(x) dx$ parametrische, stetige Linearform, $g(\cdot; \mu) \in C^0(\bar{\Omega}, \mathbb{R})$. Wähle $S_{train} \subset \mathcal{P}$ endliche Teilmenge, setze $G := \{g(\cdot; \mu) | \mu \in S_{train}\}$. Konstruiere EI-Basis $Q_m = \{q_i\}_{i=1}^M$ und EI-Punkte $T_M = \{x_i\}_{i=1}^M$ gemäß Definition 6.1. Dann ist $g(x; \mu) \approx \sum_{m=1}^M \Theta^m(\mu) q^m(x)$ also

$$f(v; \mu) \approx \sum_{m=1}^M \Theta^m(\mu) \int_\Omega q^m(x) v(x) dx =: \sum_{m=1}^M \Theta^m(\mu) f^m(v)$$

mit Komponenten $f^m(v)$ und Koeffizienten $\Theta^m(\mu)$

$$f^m(v) := \int_\Omega q^m(x) v(x) dx \quad \text{bzw.} \quad (\Theta^1(\mu), \dots, \Theta^m(\mu))^T := \underline{Q}_M^{-1} \underline{g}(\mu)$$

wobei $\underline{g}(\mu)$ durch lokale Auswertungen definiert ist

$$\underline{g}(\mu) := (g(x_1, \mu), \dots, g(x_M, \mu))^T.$$

Bemerkung:

- i) Analoge Vorgehensweise bei Ausgabefunktional oder Bilinearform möglich, anschließend Offline/Online-Zerlegung nach Kapitel 3 realisierbar.

- ii) Die Fehlerschätzer aus Kapitel 3 sind mit EI im Allgemeinen nicht rigoros, da der Interpolationsfehler zusätzlich berücksichtigt werden muss. Falls $g(\cdot, \mu) \in \text{span}(Q_M)$ (starke Annahme), so sind Schätzer weiterhin rigoros.

Kapitel 7

Allgemeine lineare Probleme

7.1 RB-Verfahren für allgemeine koerzive Probleme

Definition 7.1 (Primal-Duales detailliertes Modell):

Sei X HR, $\mathcal{P} \subset \mathbb{R}^p$, $b : X \times X \times \mathcal{P} \rightarrow \mathbb{R}$ stetige, parametrische, koerzive Bilinearform, $f, l : X \times \mathcal{P} \rightarrow \mathbb{R}$ stetige, parametrische Linearform. Dann definieren wir die primale detaillierte Lösung $u^{pr}(\mu) \in X$ durch

$$b(u^{pr}(\mu), v; \mu) = f(v; \mu) \quad \forall v \in X$$

die duale detaillierte Lösung $u^{du}(\mu) \in X$ durch

$$b(v, u^{du}(\mu); \mu) = -l(v; \mu) \quad \forall v \in X$$

und die Ausgabe

$$s(\mu) = l(u^{pr}(\mu); \mu).$$

Bemerkung:

- i) Obiges umfasst nichtsymmetrisches b
- ii) Koerzivität garantiert Existenz und Eindeutigkeit der primalen und dualen Lösung
- iii) Das duale (oder adjungierte) Problem ist zur Berechnung von $s(\mu)$ nicht notwendig, jedoch muss $u^{du}(\mu)$ im folgenden RB-Verfahren approximiert werden, um gute Ausgabeschätzungen zu erhalten.

- iv) Falls b symmetrisch, $l = f \Rightarrow u^{pr}(\mu) = -u^{du}(\mu)$, also duales Problem überflüssig.
- v) Ein Beispiel für koerzives, nichtsymmetrisches Problem: Advektion-Diffusion:
Sei $\Omega \subset \mathbb{R}^d, \underline{c} \in \mathbb{R}^d, \mu > 0$, finde u so dass

$$\nabla \cdot (\underline{c}u) - \mu \Delta u = 0 \text{ in } \Omega, \quad u = 0 \text{ auf } \partial\Omega.$$

\Rightarrow schwache Form: finde $u \in H_0^1(\Omega)$ mit

$$\int_{\Omega} u(\underline{c} \cdot \nabla v) + \int_{\Omega} \mu \nabla u \cdot \nabla v = 0 \quad \forall v \in H_0^1(\Omega)$$

- vi) Literatur: [Rov03]

D.Rovas: *Reduced Basis Output Bound Methods for parametrized PDE's*.
PhD-Thesis, MIT, Juni 2003.

Definition 7.2 (Primal-Duales RB-Modell):

Sei ein Problem gemäß Definition 7.1 gegeben. Seien $X_N^{pr}, X_N^{du} \subset X$ zwei RB-Räume mit Dimension $N^{pr}, N^{du} \in \mathbb{N}$. Zu $\mu \in \mathcal{P}$ ist die primale RB-Lösung $u_N^{pr}(\mu) \in X_N^{pr}$ definiert durch

$$b(u_N^{pr}(\mu), v; \mu) = f(v; \mu) \quad \forall v \in X_N^{pr}$$

die duale RB-Lösung $u_N^{du}(\mu) \in X_N^{du}$ durch

$$b(v, u_N^{du}(\mu); \mu) = -l(v; \mu) \quad \forall v \in X_N^{du}$$

und die RB-Ausgabe $s_N(\mu) \in \mathbb{R}$

$$s_N(\mu) = l(u_N^{pr}(\mu) - r^{pr}(u_N^{du}(\mu); \mu))$$

mit primalem Residuum $r^{pr}(\cdot; \mu) \in X'$

$$r^{pr}(v; \mu) := f(v; \mu) - b(u_N^{pr}(\mu), v; \mu)$$

und dualem Residuum $r^{du}(\cdot; \mu) \in X'$

$$r^{du}(v; \mu) := -l(v; \mu) - b(v, u_N^{du}(\mu); \mu)$$

Bemerkung:

- i) Dimensionen N^{pr}, N^{du} meist unterschiedlich. X_N^{pr} durch Snapshots von $u^{pr}(\mu)$ generiert, X_N^{du} durch Snapshots von $u^{du}(\mu)$, welche sehr verschieden sein können.

- ii) b symmetrisch, $l = f \Rightarrow r^{pr}(u^{du}(\mu); \mu) = 0$ also $s_N(\mu)$ aus Kapitel 3 reproduziert.
- iii) Existenz, Eindeutigkeit und Beschränktheit durch Daten analog zu Satz 3.4. \Rightarrow Übung.
- iv) Numerische Stabilität analog zu Satz 3.6, falls orthonormale Basen für X_N^{pr}, X_N^{du} .
- v) Falls b asymmetrisch \Rightarrow Galerkin-Projektion keine orthogonale Projektion mehr, aber Galerkin-Orthogonalität gilt weiterhin:

$$b(u^{pr}(\mu) - u_N^{pr}(\mu), v; \mu) = 0 \quad v \in X_N^{pr}$$

und analog für das duale Problem.

- vi) Residuen sind wieder „rechte Seiten“ der Fehlergleichung

$$b(u^{pr}(\mu) - u_N^{pr}(\mu), v; \mu) = r^{pr}(v; \mu)$$

und analog für das duale Problem.

- vii) „Reproduktion“ von Lösungen gemäß Korollar 3.8 gilt weiterhin.
- viii) Offline/Online-Zerlegung folgt ähnlich zu Abschnitt 3.2 im Fall separabler Parameterabhängigkeit.

Satz 7.3 (Beschränktheit durch Bestapproximation):

- i) Für den Ausgabefehler gilt

$$\begin{aligned} |s(\mu) - s_N(\mu)| &= |b(u^{pr}(\mu) - u_N^{pr}(\mu), u^{du}(\mu) - u_N^{du}(\mu))| \\ &\leq \gamma(\mu) \|u^{pr}(\mu) - u_N^{pr}(\mu)\| \|u^{du}(\mu) - u_N^{du}(\mu)\| \end{aligned}$$

- ii) Für den Fehler in primaler/dualer Lösung gilt

$$\begin{aligned} \|u^{pr}(\mu) - u_N^{pr}(\mu)\| &\leq \frac{\gamma(\mu)}{\alpha(\mu)} \inf_{v \in X_N^{pr}} \|u^{pr}(\mu) - v\| \\ \|u^{du}(\mu) - u_N^{du}(\mu)\| &\leq \frac{\gamma(\mu)}{\alpha(\mu)} \inf_{v \in X_N^{du}} \|u^{du}(\mu) - v\| \end{aligned}$$

Beweis:

i)

$$\begin{aligned}
|s(\mu) - s_N(\mu)| &= |l(u^{pr}) - l(u_N^{pr}) + r^{pr}(u_N^{du})| \\
&= |l(u^{pr} - u_N^{pr}) + f(u_N^{du}) - b(u_N^{pr}, u_N^{du})| \\
&= | -b(u^{pr} - u_N^{pr}, u^{du}) + b(u^{pr}, u_N^{du}) - b(u_N^{pr}, u_N^{du}) | \\
&= | -b(u^{pr} - u_N^{pr}, u^{du}) + b(u^{pr} - u_N^{pr}, u_N^{du}) | \\
&= | -b(u^{pr} - u_N^{pr}, u^{du} - u_N^{du}) | \\
&\leq \gamma \|u^{pr} - u_N^{pr}\| \|u^{du} - u_N^{du}\|.
\end{aligned} \tag{7.1}$$

ii) „Lemma von Cea“

$$\begin{aligned}
\alpha \|u^{pr} - u_N^{pr}\|^2 &\leq b(u^{pr} - u_N^{pr}, u^{pr} - u_N^{pr}) = b(u^{pr} - u_N^{pr}, u^{pr} - v) \\
&\leq \gamma \|u^{pr} - u_N^{pr}\| \|u^{pr} - v\| \quad \forall v \in X_N^{pr}
\end{aligned}$$

Kürzen und Infimum liefert Behauptung. Analog für duales Problem. □**Satz 7.4** (A-Posteriori-Fehlerschätzer):

Seien $v_r^{pr}(\mu), v_r^{du}(\mu) \in X$ die Riesz-Repräsentanten der Residuen und $\alpha_{LB}(\mu)$ eine untere Schranke für die Koerzivitätskonstante von $b(\cdot, \cdot; \mu)$. Dann gilt

i) für den primalen Fehler $e^{pr}(\mu) := u^{pr}(\mu) - u_N^{pr}(\mu)$

$$\|e^{pr}(\mu)\| \leq \Delta_N^{pr}(\mu) := \frac{\|v_r^{pr}(\mu)\|}{\alpha_{LB}(\mu)},$$

ii) für den dualen Fehler $e^{du}(\mu) := u^{du}(\mu) - u_N^{du}(\mu)$

$$\|e^{du}(\mu)\| \leq \Delta_N^{du}(\mu) := \frac{\|v_r^{du}(\mu)\|}{\alpha_{LB}(\mu)},$$

iii) für den Ausgabefehler

$$|s(\mu) - s_N(\mu)| \leq \Delta_N^s(\mu) := \frac{\|v_r^{pr}\| \cdot \|v_r^{du}\|}{\alpha_{LB}(\mu)}.$$

Beweis:

i)

$$\begin{aligned}
\alpha_{LB}(\mu) \|e^{pr}(\mu)\|^2 &\leq b(e^{pr}, e^{pr}; \mu) = r^{pr}(e^{pr}) \\
&= (v_r^{pr}, e^{pr}) \leq \|v_r^{pr}\| \|e^{pr}\|
\end{aligned}$$

 \Rightarrow Kürzen liefert die Behauptung.

ii) analog zu i)

iii)

$$\begin{aligned} |s(\mu) - s_N(\mu)| &\stackrel{(7.1)}{=} |b(e^{pr}, e^{du})| = |r^{pr}(e^{du})| = |(v_r^{pr}, e^{du})| \\ &\leq \|v_r^{pr}\| \|e^{du}\| \stackrel{ii)}{\leq} \|v_r^{pr}\| \frac{\|v_r^{du}\|}{\alpha_{LB}(\mu)}. \end{aligned}$$

□

Bemerkung:

- i) wieder „quadratischer“ Effekt in der Ausgabeschranke. Dies liefert ein Kriterium zum Verhältnis von N^{pr} und N^{du} : Wähle diese so, das $\|v_r^{pr}(\mu)\|$ und $\|v_r^{du}(\mu)\|$ etwa gleich groß.
- ii) Offline/Online Zerlegung der Residuen und Fehlerschätzer analog zu Kap. 3.
- iii) Falls $\alpha(\mu)$ und somit $\alpha_{LB}(\mu)$ sehr klein, so sind Abschätzungen sehr grob. Eventuell sind stärkere Aussagen mit anderem Stabilitätsbegriff möglich: Inf-sup-Stabilität.

7.2 Inf-sup-Stabilität

Definition 7.5 (Inf-sup-Konstante/Stabilität):

Seien X_1, X_2 HR, $b : X_1 \times X_2 \rightarrow \mathbb{R}$ bilinear und stetig. Wir definieren die Inf-sup-Konstante

$$\beta := \inf_{u \neq 0} \sup_{v \neq 0} \frac{b(u, v)}{\|u\| \|v\|}$$

Falls $\beta > 0$ ist b inf-sup-stabil.

Bemerkung:

Es gilt immer $\beta \geq 0$.

Lemma 7.6 (Eigenschaften von β/B):

Sei $b(u, v)$ stetige Bilinearform und $b(u, v) = (Bu, v)$ mit geeignetem $B \in L(X_1, X_2)$ nach Satz 2.8. Dann gilt:

i) B ist „suprimierender Operator“

$$\sup_{v \in X_2} \frac{b(u, v)}{\|v\|_{X_2}} = \frac{b(u, Bu)}{\|Bu\|_{X_2}} \quad \forall u \in X_1 \setminus \{0\}$$

ii) $\beta = \inf_{u \neq 0} \frac{\|Bu\|}{\|u\|}$ also insbesondere $\beta \geq 0$

iii) $\gamma = \sup_{u \neq 0} \frac{\|Bu\|}{\|u\|}$

iv) $\forall u \in X_1 \exists v \in X_2$, so dass

$$\beta \|u\| \|v\| \leq b(u, v).$$

Beweis:

i) “ \geq “: klar wegen Supremum. “ \leq “: Sei $v \in X_2 \setminus \{0\}$, $u \in X_1 \setminus \{0\}$

$$\begin{aligned} \frac{b(u, v)}{\|v\|} &= \frac{(Bu, v)}{\|v\|} \leq \frac{\|Bu\| \|v\|}{\|v\|} = \frac{\|Bu\| \|Bu\|}{\|Bu\|} \\ &= \frac{(Bu, Bu)}{\|Bu\|} = \frac{b(u, Bu)}{\|Bu\|} \end{aligned}$$

Also gilt dies auch für Supremum.

ii)

$$\begin{aligned} \beta &= \inf_{u \neq 0} \sup_{v \neq 0} \frac{b(u, v)}{\|u\| \|v\|} \stackrel{i)}{=} \inf_{u \neq 0} \frac{b(u, Bu)}{\|u\| \|Bu\|} = \inf_{u \neq 0} \frac{(Bu, Bu)}{\|u\| \|Bu\|} \\ &= \inf_{u \neq 0} \frac{\|Bu\| \|Bu\|}{\|u\| \|Bu\|} = \inf_{u \neq 0} \frac{\|Bu\|}{\|u\|}. \end{aligned}$$

iii) analog zu ii)

iv) Nach i) ist $\beta = \inf_{u \neq 0} \frac{b(u, Bu)}{\|u\| \|Bu\|}$. Also für alle $u \in X_1$:

$$\beta \leq \frac{b(u, Bu)}{\|u\| \|Bu\|} \Rightarrow \beta \|u\| \|Bu\| \leq b(u, Bu)$$

mit $v := Bu$ folgt die Behauptung, für $u = 0$ gilt Beh. sowieso. □

Satz 7.7 (Eigenwertproblem für β, γ):

Seien X_1, X_2 HR, $b : X_1 \times X_2$ stetige Bilinearform mit $b(u, v) = (Bu, v)$ für $B \in K(X_1, X_2)$, $B^* \in K(X_2, X_1)$ und $\sigma := \sigma(B^*B)$ das Spektrum von B^*B . Dann gilt

i) (EWP- β): $\forall \lambda \in \sigma \exists u \in X_1$:

$$(Bu, Bv) = \lambda(u, v) \quad \forall v \in X_1$$

ii) $\beta^2 = \inf \sigma$

iii) $\gamma^2 = \max \sigma$

Beweis:

i) Seien (u, λ) EV, EW von $B^*B \Rightarrow (B^*Bu, v) = (\lambda u, v) \quad \forall v \in X_1 \Rightarrow (Bu, Bv) = \lambda(u, v)$

ii) und iii) siehe Übung. □

Korollar 7.8 (Inf-sup-Stabilität im Endlichdimensionalen):

Seien $X_1 = \mathbb{R}^m, X_2 = \mathbb{R}^n, B \in \mathbb{R}^{n \times m}, m \leq n$ und $b(u, v) = (Bu, v)$. Seien $\{\sigma_i\}_{i=1}^m$ Singulärwerte von B . Dann

i) $\beta = \min_{i=1, \dots, m} \sigma_i$

ii) b inf-sup-stabil $\Leftrightarrow B$ hat vollen Spaltenrang.

Beweis:

i) b stetig und B, B^T kompakt, da endlichdimensionale Bilder. Für $\sigma(B^T B) = \{\sigma_i^2\}$ gilt mit Satz 7.7: $\beta^2 = \inf \sigma(B^T B) = \min \sigma_i^2$. Da $\beta \geq 0, \sigma_i \geq 0 \Rightarrow \beta = \min_i \sigma_i$.

ii) b inf-sup-stabil $\Leftrightarrow \beta > 0 \Leftrightarrow \forall i = 1, \dots, m : \sigma_i > 0 \Leftrightarrow B$ hat Rang $m \Leftrightarrow B$ hat vollen Spaltenrang □

Lemma 7.9 (Beziehung zwischen Inf-sup-Stabilität und Koerzivität):

Sei X HR, $b : X \times X \rightarrow \mathbb{R}$ stetige Bilinearform

i) Falls b koerziv $\Rightarrow b$ inf-sup-stabil mit

$$\alpha \leq \beta.$$

ii) Sei $b(u, v) = (Bu, v)$ mit $B \in K(X)$, b symmetrisch, koerziv $\Rightarrow b$ inf-sup-stabil mit

$$\alpha = \beta.$$

Beweis:

$$i) \beta = \inf_u \sup_v \frac{b(u,v)}{\|u\| \|v\|} \geq \inf_u \frac{b(u,u)}{\|u\|^2} = \alpha > 0.$$

ii) Sei (u, λ) EV,EW von $(EWP-\alpha)$ aus Satz 2.9, d.h. $b_s(u, v) = \lambda(u, v) \forall v \in X$. Dann ist

$$\begin{aligned} (Bu, Bv) &\stackrel{Def}{=} b(u, Bv) \stackrel{b \text{ sym}}{=} b_s(u, Bv) \stackrel{(EWP-\alpha)}{=} \lambda(u, Bv) \\ &= \lambda(Bv, u) = \lambda b(v, u) \stackrel{sym}{=} \lambda b_s(u, v) \\ &= \lambda^2(u, v) \end{aligned}$$

$\Rightarrow \lambda^2$ ist EW von $(EWP-\beta)$ aus Satz 7.7. Nach Satz 2.9 ist α^2 EW von $(EWP-\beta)$ also

$$\begin{aligned} \beta^2 &= \inf \sigma(B^*B) \leq \alpha^2 \\ \stackrel{\text{mit i)}}{\Rightarrow} \beta^2 &= \alpha^2. \end{aligned}$$

Da $\alpha, \beta > 0 \Rightarrow \beta = \alpha$.

□

7.3 RB-Verfahren für inf-sup-stabile Probleme

Inf-sup-Stabilität garantiert Wohlgestelltheit von Gleichungssystemen im Endlichdimensionalen:

Lemma 7.10 (Existenz und Eindeutigkeit und Stabilität):

Sei $b : X_1 \times X_2 \rightarrow \mathbb{R}$ inf-sup-stabil mit Inf-sup-Konstante β und $\dim X_1 = \dim X_2 < \infty$. Dann hat für $f \in X_2'$

$$b(u, v) = f(v) \quad \forall v \in X_2$$

genau eine Lösung $u \in X_1$, diese erfüllt $\|u\| \leq \frac{\|f\|}{\beta}$.

Beweis: Sei $b(u, v) = (Bu, v)$. Dann ist B injektiv, denn sonst ex. $u \in X_1 \setminus \{0\}$ mit

$$0 = Bu \Rightarrow 0 = (Bu, v) = b(u, v) \forall v \in X_2.$$

Nach Lemma 7.6 iv) gäbe es dann $v \in X_2 \setminus \{0\}$ so dass

$$0 < \beta \|u\| \|v\| \leq b(u, v) = (Bu, v) = 0 \quad \zeta$$

Da $\dim X_1 = \dim X_2$ ist B auch surjektiv also invertierbar. Sei $v_f \in X_2$ Riesz Repräsentant von f , dann sind äquivalent

$$\begin{aligned} (Bu, v) &= b(u, v) = f(v) = (v_f, v) \quad \forall v \in X_2 \\ \Leftrightarrow u &= B^{-1}v_f \end{aligned}$$

Letztes ist offensichtlich eindeutig lösbar.

$$\begin{aligned} \beta \stackrel{7.6ii)}{=} \inf \frac{\|Bu\|}{\|u\|} &= \inf \frac{\|v\|}{\|B^{-1}v\|} = \left(\sup_v \frac{\|B^{-1}v\|}{\|v\|} \right)^{-1} = \|B^{-1}\|^{-1} \\ \Rightarrow \|u\| &\leq \|B^{-1}\| \|v_f\| = \frac{\|v_f\|}{\beta} = \frac{\|f\|}{\beta}. \end{aligned}$$

□

Bemerkung:

Insbesondere können wir das RB-Verfahren aus Definition 7.1 und 7.2 unverändert übernehmen, indem wir statt Koerzivität Inf-sup-Stabilität fordern. Die folgenden Aussagen gelten für koerzive und nicht koerzive aber inf-sup-stabile Probleme.

Satz 7.11 (Beschränktheit durch Bestapproximation):

Sei ein Problem gemäß Definition 7.1 und 7.2 gegeben und $\beta_N^{pr}(\mu), \beta_N^{du}(\mu) \geq \beta_0 > 0$ für alle N^{pr}, N^{du} mit $\beta_N^{pr}, \beta_N^{du}$ Inf-sup-Konstanten von $b^{pr}(u, v; \mu) := b(u, v; \mu)$ auf $X_N^{pr} \times X_N^{pr}$ bzw. $b^{du}(u, v; \mu) := b(v, u; \mu)$ auf $X_N^{du} \times X_N^{du}$ und $\gamma(\mu)$ Stetigkeitskonstante von $b(\cdot, \cdot; \mu)$ auf $X \times X$. Dann gilt

$$\|u^{pr}(\mu) - u_N^{pr}(\mu)\| \leq \left(1 + \frac{\gamma(\mu)}{\beta_0}\right) \inf_{v \in X_N^{pr}} \|u^{pr}(\mu) - v\|$$

analog für duale lösung und Ausgabe nach 7.3i).

Beweis: Sei $b(u, v) = (Bu, v)$ für $B \in L(X_N^{pr})$. Für $v \in X_N^{pr}$ gilt

$$\begin{aligned} \beta_N^{pr}(\mu) \|v - u_N^{pr}\| \underbrace{\|B(v - u_N^{pr})\|}_{\in X_N^{pr}} &\stackrel{7.6ii)}{\leq} b(v - u_N^{pr}, \underbrace{B(v - u_N^{pr})}_{\in X_N^{pr}}) \\ &= b(v - u_N^{pr} + (u_N^{pr} - u^{pr}), B(v - u_N^{pr})) \\ &= b(v - u^{pr}, B(v - u_N^{pr})) \\ &\leq \gamma \|v - u^{pr}\| \underbrace{\|B(v - u_N^{pr})\|}_{\in X_N^{pr}} \\ \Rightarrow \|v - u_N^{pr}\| &\leq \frac{\gamma}{\beta_N^{pr}} \|v - u^{pr}\| \end{aligned}$$

$$\begin{aligned}
\|u^{pr} - u_N^{pr}\| &= \inf_{v \in X_N^{pr}} \|u^{pr} - v + v - u_N^{pr}\| \\
&\leq \inf_{v \in X_N^{pr}} (\|u^{pr} - v\| + \|v - u_N^{pr}\|) \\
&\leq \left(1 + \frac{\gamma}{\beta_0}\right) \inf_{v \in X_N^{pr}} \|v - u_N^{pr}\|.
\end{aligned}$$

Für duales Problem analog. □

Bemerkung:

- i) falls b koerziv und $(1 + \frac{\gamma}{\beta_0}) < \frac{\gamma}{\beta}$ ist Satz 7.11 eine schärfere Aussage als Satz 7.3.
- ii) Für koerzive Probleme gilt für untere Schranke $\alpha_{LB}(\mu)$ an $\alpha(\mu)$ und $\alpha_N(\mu) = \inf_{v \in X_N} \frac{b(v,v)}{\|v\|^2}$:

$$\alpha_{LB}(\mu) \leq \alpha(\mu) \leq \dots \leq \alpha_{N+1}(\mu) \leq \alpha_N(\mu)$$

da sich Koerzivität auf Teilräume „vererbt“. Für inf-sup-stabile Probleme gilt dies nicht, $\beta(\mu), \beta_N(\mu), \beta_{N+1}(\mu)$ sind vollkommen unabhängig. \Rightarrow Übung.

Satz 7.12 (A-Posteriori-Fehlerschätzer):

Sei ein Problem gemäß 7.1 und 7.2 gegeben mit $\beta_{LB}(\mu) > 0$ untere Schranke für die Inf-sup-Konstante $\beta(\mu)$ von $b^{pr}(u, v; \mu) = b(u, v; \mu)$ und $b^{du}(u, v; \mu) = b(v, u; \mu)$ auf $X \times X$. Dann gelten

- i) $\|u^{pr}(\mu) - u_N^{pr}(\mu)\| \leq \Delta_N^{pr}(\mu) := \frac{\|v_r^{pr}(\mu)\|}{\beta_{LB}(\mu)}$
- ii) $\|u^{du}(\mu) - u_N^{du}(\mu)\| \leq \Delta_N^{du}(\mu) := \frac{\|v_r^{du}(\mu)\|}{\beta_{LB}(\mu)}$
- iii) $|s(\mu) - s_N(\mu)| \leq \Delta_N^s(\mu) := \frac{\|v_r^{pr}\| \|v_r^{du}\|}{\beta_{LB}(\mu)}$

Beweis:

- i) Nach Definition gilt für bestimmtes $v \neq 0$:

$$\beta_{LB}(\mu) \|e^{pr}\| \|v\| \leq b(e^{pr}, v; \mu) = r^{pr}(v) = (v_r^{pr}, v) \leq \|v_r^{pr}\| \|v\|$$

Also

$$\|e^{pr}\| \leq \frac{\|v_r^{pr}\|}{\beta_{LB}(\mu)}.$$

ii) analog zu i)

iii) analog zu Satz 7.4 iii)

□

Bemerkung: i) Falls $\beta_N^{pr}(\mu) \rightarrow \beta^{pr}(\mu)$ gleichmäßig in \mathcal{P} für $N^{pr} \rightarrow \infty$, dann besteht folgende Möglichkeit zur Konstruktion von $\beta_{LB}(\mu)$: Nach Satz 7.7 bestimme λ_{min} kleinster EW von (EWP- β) und damit $\beta^{pr}(\mu) := \sqrt{\lambda_{min}}$. Wähle $\sigma \in (0, 1)$ und setze $\beta_{LB}(\mu) := \sigma \cdot \beta^{pr}(\mu)$. Dann ist für genügend großes N^{pr} $\beta_{LB}(\mu) \leq \beta_N^{pr}(\mu)$.

ii) Als alternative zu i) oder Satz 3.15 ist die „Successive Constraints Method“ [HRSP07] ein allgemeines Verfahren, welches untere Schranken α_{LB}, β_{LB} durch Lösen eines (kleinen) Optimierungsproblems bestimmen kann. Dies ist auch effizient offline/online zerlegbar, siehe Kap. 8.

iii) Um „Stabilität“ zu steigern, d.h. $\beta_N^{pr}, \beta_N^{du}$ zu vergrößern, ist statt Galerkin-Projektion ein Petrov-Galerkin Ansatz möglich:

Seien $X_{N,1}^{pr}, X_{N,2}^{pr}$ unterschiedliche Ansatz- und Testräume. Suche $u_N^{pr} \in X_{N,1}^{pr}$

$$b(u_N^{pr}, v) = f(v) \quad \forall v \in X_{N,2}^{pr}$$

Hierbei ist $X_{N,1}^{pr}$ für gute Approximationen und $X_{N,2}^{pr}$ für gute Stabilität konstruiert.

Kapitel 8

Weiterführende Aspekte

8.1 Successive Constraints Method

Für a-posteriori Fehlerschranken sind schnell berechenbare Schranken für Stabilitätskonstanten erforderlich. Diese sind mit dem folgenden Verfahren berechenbar, siehe [HRSP07].

Definition 8.1 (Successive Constraints Method (SCM)):

Sei $b(u, v; \mu) = \sum_{q=1}^Q \theta_b^q(\mu) b^q(u, v)$ auf $X \times X \times P$ parametrisch koerziv mit Koerzivitätskonstante $\alpha(\mu)$ und \mathcal{P} kompakt. Seien $C, D \subset \mathcal{P}$ endliche Teilmengen und $M_\alpha, M_+ \in \mathbb{N}$. Definiere

$$Y := \left\{ y = (y_1, \dots, y_Q) \in \mathbb{R}^Q \mid \exists u \in X \text{ mit } y_q = \frac{b^q(u, u)}{\|u\|^2}, 1 \leq q \leq Q \right\}$$

Wir definieren eine Zielfunktion $J : \mathcal{P} \times \mathbb{R}^Q \rightarrow \mathbb{R}$

$$J(\mu, y) := \sum_{q=1}^Q \theta_b^q(\mu) y_q,$$

und ein Polytop durch

$$\sigma_q^- := \inf_u \frac{b^q(u, u)}{\|u\|^2}, \quad \sigma_q^+ := \sup_u \frac{b^q(u, u)}{\|u\|^2},$$

$$B_Q := \prod_{q=1}^Q [\sigma_q^-, \sigma_q^+] \subset \mathbb{R}^Q.$$

Für $M \in \mathbb{N}, \mu \in \mathcal{P}$ definiere $P_M(\mu, C) \subset C$ durch

$$P_M(\mu, C) := \begin{cases} M - \text{nächsten Nachbarn von } \mu \in C & \text{falls } 1 \leq M < |C| \\ C & \text{falls } |C| \leq M \\ \emptyset & \text{falls } M = 0. \end{cases}$$

Wir definieren hiermit für $\mu \in \mathcal{P}$:

$$\begin{aligned} Y_{LB}(\mu) &:= \{y \in B_Q \mid \\ &\quad J(\mu', y) \geq \alpha(\mu') \quad \forall \mu' \in P_{M_\alpha}(\mu, C) \text{ und} \\ &\quad J(\mu', y) \geq 0 \quad \forall \mu' \in P_{M_+}(\mu, D)\} \\ Y_{UB} &:= \{y^*(\mu') \mid \mu' \in C\} \quad \text{mit} \quad y^*(\mu') := \arg \min_{y \in Y} J(\mu', y) \end{aligned}$$

und hiermit die Schranken

$$\alpha_{LB}(\mu) := \min_{y \in Y_{LB}(\mu)} J(\mu, y), \quad (8.1)$$

$$\alpha_{UB}(\mu) := \min_{y \in Y_{UB}} J(\mu, y).$$

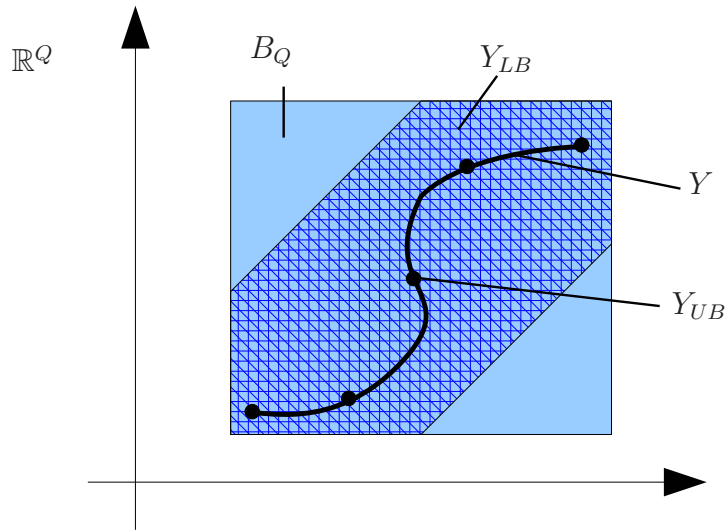


Abbildung 8.1: Mengen der Successive Constraint Method

Satz 8.2 (α -Approximation mit SCM):

Es gilt für alle $\mu \in \mathcal{P}$

$$\alpha_{LB}(\mu) \leq \alpha(\mu) \leq \alpha_{UB}(\mu)$$

Beweis:

Zunächst sehen wir, dass

$$\alpha(\mu) = \inf_u \frac{\sum_q \theta_b^q(\mu) b^q(u, u)}{\|u\|^2} = \min_{y \in Y} J(\mu, y).$$

Es gilt weiter

$$Y_{UB} \subset Y \subset Y_{LB}(\mu).$$

Erste Inklusion: $y \in Y_{UB} \Rightarrow y = y^*(\mu') = \arg \min_{y \in Y} J(\mu', y)$ für ein $\mu' \in C$, also $y \in Y$. Zweite Inklusion: Für $y \in Y$ gilt $y \in B_Q$ und

$$\alpha(\mu) = \min_{\bar{y} \in Y} J(\mu', \bar{y}) \leq J(\mu', y) \quad \forall \mu' \in C$$

und analog die Positivitätsbedingung. Also gilt

$$\underbrace{\min_{y \in Y_{LB}(\mu)} J(\mu, y)}_{=\alpha_{LB}(\mu)} \leq \underbrace{\min_{y \in Y} J(\mu, y)}_{=\alpha(\mu)} \leq \underbrace{\min_{y \in Y_{UB}} J(\mu, y)}_{=\alpha_{UB}(\mu)}.$$

□

Bemerkung:

- i) Für festes μ ist J also insbesondere linear in y also (8.1) ein kleines, schnell lösbares, lineares Optimierungsproblem.
- ii) C, D durch ein Greedy-Verfahren konstruierbar
- iii) Ähnlich sind Schranken für β_N und γ berechenbar
- iv) SCM ist immer schärfer als der $\min \theta$ -Ansatz aus Lemma 3.15.

8.2 Nichtlineare Probleme

Behandlung als Multilinearform:

Beispiel viskose Burgers-Gleichung, siehe [VPP03]

$$\partial_x(u^2) - \varepsilon \partial_{xx} u = g \quad \text{auf } \Omega = (0, 1), \quad u(0) = u(1) = 0.$$

Schwache Form: suche $u \in X := H_0^1(\Omega)$ mit

$$\int_{\Omega} \partial_x(u^2)v - \varepsilon(\partial_{xx} u)v = \int_{\Omega} gv \quad \forall v \in X$$

Partielle Integration liefert

$$-\int_{\Omega} u^2 \partial_x v + \int_{\Omega} \varepsilon \partial_x u \partial_x v = \int_{\Omega} g v$$

Setze Multilinearformen $a(u, v, w) := -\int_{\Omega} uv \partial_x w$, $b(u, v) := \int_{\Omega} \varepsilon \partial_x u \partial_x v$, $f(v) := \int_{\Omega} g v$. Suche also $u \in X$ mit

$$\underbrace{a(u, u, \cdot) + b(u, \cdot) - f(\cdot)}_{:=F(u) \in X'} = 0.$$

Löse $F(u) = 0$ durch Newton-Verfahren. Hierzu Richtungs-Ableitungen bequem durch Multilinearformen darstellbar: Für alle $u, h \in X$ gilt:

$$DF|_u(h) = \lim_{\delta \rightarrow 0} \frac{F(u + \delta h) - F(u)}{\delta}.$$

$$\begin{aligned} F(u + \delta h) - F(u) &= a(u + \delta h, u + \delta h, \cdot) - a(u, u, \cdot) \\ &\quad + b(u + \delta h, \cdot) - b(u, \cdot) - f(\cdot) + f(\cdot) \\ &= 2\delta a(u, h, \cdot) + \delta^2 a(h, h, \cdot) + \delta b(h, \cdot). \end{aligned}$$

Also

$$DF|_u(h) = 2a(u, h, \cdot) + b(h, \cdot).$$

Newton-Schleife:

- Wähle $u^0 \in X$
- Wiederhole
bestimme h^k Lösung von $DF|_{u^k}(h^k) = -F(u^k)$ also

$$2a(u^k, h^k, v) + b(h^k, v) = -a(u^k, u^k, v) - b(u^k, v) + f(v) \quad \forall v \in X$$

$$\text{Setze } u^{k+1} := u^k + h^k$$

- Bis $\|u^{k+1} - u^k\| < \varepsilon_{tol}$.

RB-Verfahren ist aus Galerkin-Projektion dieser Schleife sofort ersichtlich.

Bemerkung:

- i) Neben Offline-Matrizen nun auch Offline-Tensoren erforderlich: $A_{ijk}^q := a^q(\varphi_i, \varphi_j, \varphi_k)$. Systemassemblierung für h^k in der Online-Phase offensichtlich Komplexität $O(N^m)$ mit m Anzahl der Argumente der führenden Multilinearform. Verfahren also nur für kleine $m = 2, 3, 4$ praktikabel.
- ii) Allgemeinere Nichtlinearitäten: Direkt mittels Empirischer Interpolation.

Ausblick auf das Wintersemester WS 2010/2011:

Vorlesung "Reduzierte Basis Methoden II"

- RB für instationäre Probleme
- RB für Optimierung
- Basisgenerierung
- ...

→ BSc- und Diplomarbeiten.

Literaturverzeichnis

- [Alt92] H.W. Alt. *Lineare Funktionalanalysis*. Springer, 1992.
- [Ant05] A.C. Antoulas. *Approximation of Large-Scale Dynamical Systems*. SIAM Publications, Philadelphia, PA, 2005.
- [BCD⁺10] P. Binev, A. Cohen, W. Dahmen, R. DeVore, G. Petrova, and P. Wojtaszczyk. Convergence rates for greedy algorithms in reduced basis methods. *IGPM Report, RWTH Aachen*, 310, 2010.
- [BMNP04] M. Barrault, Y. Maday, N.C. Nguyen, and A.T. Patera. An 'empirical interpolation' method: application to efficient reduced-basis discretization of partial differential equations. *C. R. Math. Acad. Sci. Paris Series I*, 339:667–672, 2004.
- [Bra03] D. Braess. *Finite Elemente*. Springer, 2003.
- [FR83] J.P. Fink and W.C. Rheinboldt. On the error behaviour of the reduced basis technique for nonlinear finite element approximations. *ZAMM*, 63:21–28, 1983.
- [HRSP07] D.B.P. Huynh, G. Rozza, S. Sen, and A.T. Patera. A successive constraint linear optimization method for lower bounds of parametric coercivity and inf-sup stability constants. *C. R. Math. Acad. Sci. Paris Series I*, 345:473–478, 2007.
- [MNPP07] Y. Maday, N.C. Nguyen, A.T. Patera, and G.S.H. Pau. A general, multi-purpose interpolation procedure: the magic points. Technical Report Preprint RO7037, Laboratoire Jacques-Louis-Lions, Université Pierre et Marie Curie, Paris, 2007.
- [MPT02] Y. Maday, A.T. Patera, and G. Turinici. a priori convergence theory for reduced-basis approximations of single-parameter symmetric coercive elliptic partial differential equations. *C.R. Acad. Sci. Paris, Ser. I*, 335:289–294, 2002.

- [PR07] A.T. Patera and G. Rozza. *Reduced Basis Approximation and a Posteriori Error Estimation for Parametrized Partial Differential Equations*. MIT, 2007. Version 1.0, Copyright MIT 2006-2007, to appear in (tentative rubric) MIT Pappalardo Graduate Monographs in Mechanical Engineering.
- [Rov03] D.V. Rovas. *Reduced-Basis Output Bound Methods for Parametrized Partial Differential Equations*. PhD thesis, MIT, Cambridge, MA, 2003.
- [VPP03] K. Veroy, C. Prud'homme, and A.T. Patera. Reduced-basis approximation of the viscous Burgers equation: rigorous a posteriori error bounds. *C. R. Math. Acad. Sci. Paris Series I*, 337:619–624, 2003.

Bernard Haasdonk
University of Stuttgart,
Institute of Applied Analysis and Numerical Simulation,
Pfaffenwaldring 57
70569 Stuttgart
Germany
E-Mail: haasdonk@mathematik.uni-stuttgart.de

Erschienene Preprints ab Nummer 2011/001

Komplette Liste: <http://preprints.ians.uni-stuttgart.de>

- 2011/001 *Kutter, M., Eck, Ch.:* On the solvability of a two scale model for liquid phase epitaxy with elasticity
- 2011/002 *Lalegname, A., Sändig, A.-M.:* Wave-crack interaction in finite elastic bodies
- 2011/003 *Dihlmann, M., Drohmann, M., Haasdonk, B.:* Model reduction of parametrized evolution problems using the reduced basis method with adaptive time partitioning
- 2011/004 *Haasdonk, B.:* Reduzierte-Basis-Methoden