

Untersuchung aktiver Unterräume und deren Anwendung

Hendrik Kleikamp

Geboren am 28. März 1997 in Ahlen

31. Juli 2018

Bachelorarbeit Mathematik

Betreuer: Prof. Dr. Jochen Garcke

Zweitgutachter: Prof. Dr. Michael Griebel

INSTITUT FÜR NUMERISCHE SIMULATION

MATHEMATISCH-NATURWISSENSCHAFTLICHE FAKULTÄT DER
RHEINISCHEN FRIEDRICH-WILHELMS-UNIVERSITÄT BONN

Inhaltsverzeichnis

1	Einleitung	1
2	Theoretische Berechnung aktiver Unterräume	5
2.1	Grundlegende Notation und Annahmen	5
2.2	Theoretischer Ansatz	7
2.2.1	Unzentrierte Kovarianzmatrix der ersten partiellen Ableitungen . . .	7
2.2.2	Bestimmung aktiver und inaktiver Richtungen	9
2.2.3	Beispiel	11
2.2.4	Zusammenfassung	13
3	Umsetzung des Verfahrens	15
3.1	Approximation der Matrix C	15
3.1.1	Abstand der Eigenwerte	17
3.1.2	Abstand der Unterräume	18
3.2	Alternativen zur Verwendung von Gradienten	21
3.2.1	Approximation von Gradienten	22
3.2.2	Lokale lineare Modelle	22
3.2.3	Globale lineare Modelle	24
3.3	Eigenwertzerlegung der Matrix C	25
3.4	Kombinierter aktiver Unterraum zu mehreren Funktionen	26
4	Berechnung aktiver Unterräume in praktischen Anwendungen	31
4.1	Simulation des Krankheitsverlaufs nach einer Infektion mit dem HI-Virus .	31
4.1.1	Berechnung der Zahl der T-Helferzellen	32
4.1.2	Aktive Unterräume zu einem festen Zeitpunkt	33
4.1.3	Sufficient summary plots	35
4.1.4	Globale Approximation der T-Helferzellen-Population	39
4.1.5	Kombinierter aktiver Unterraum für mehrere Zeitpunkte	41
4.2	Optimierung der Profilgeometrie von Flugzeugtragflächen	43
4.2.1	Modellierung des Flügelquerschnitts	43
4.2.2	Simulation der Strömung um den Flügel	46
4.2.3	Aktive Unterräume in Auftriebs- und Reibungskoeffizienten	47
5	Anwendung auf Optimierungsprobleme	51
5.1	Aktive Unterräume in Parameterstudien	51
5.2	Aktive Unterräume in Optimierungsaufgaben	56
5.2.1	Konstruktion einer Approximation der Funktion f	56
5.2.2	Optimierung der Flügelgeometrie unter Nebenbedingungen	59
6	Zusammenfassung und Ausblick	63
	Abbildungsverzeichnis	65
	Literaturverzeichnis	67

Kapitel 1

Einleitung

Die Untersuchung komplexer Modelle und Simulationen aus unterschiedlichen Bereichen der Forschung erfordert oftmals aufwändige Parameterstudien von Funktionen mit hochdimensionalem Definitionsbereich. Modelle aus Physik, Biologie, Medizin, Aerodynamik oder anderen Anwendungsgebieten führen beispielsweise zu Systemen partieller oder gewöhnlicher Differentialgleichungen, deren approximative Lösung bereits einen hohen Rechenaufwand benötigt. Viele der betrachteten Parameterstudien leiden unter dem sogenannten „Fluch der Dimensionalität“ [Bel61]. In einer Dimension sind effiziente Algorithmen mit zufriedenstellenden Konvergenzeigenschaften bekannt. Allerdings steigt die Anzahl der Funktionsauswertungen in den hochdimensionalen Verallgemeinerungen dieser Verfahren exponentiell mit wachsender Dimension an, sofern die Konvergenz erhalten bleiben soll. Insbesondere, wenn die Berechnung eines einzelnen Funktionswertes aufwändig ist, sind Optimierung, Integration oder Regression in hohen Dimensionen kaum durchführbar.

Um die genannten Parameterstudien trotzdem zu ermöglichen, wird versucht, die Verfahren auf einige wenige wichtige Parameter zu beschränken, also die Dimension des Parameterraums zu reduzieren. Dieses Vorgehen ist für viele Probleme nicht ausreichend, da die Modelle immer noch von vielen Parametern abhängen und nur wenige Parameter komplett entfernt werden können, ohne die Ergebnisse zu verfälschen. Einen allgemeineren Ansatz stellen *Aktive Unterräume* dar, die wichtige *Richtungen* im Parameterraum, das heißt Linearkombinationen von Parametern, identifizieren. Entlang dieser Richtungen verändert sich die betrachtete Funktion signifikant, während die Funktion entlang der entsprechenden orthogonalen Richtungen nur geringe Veränderungen in den Funktionswerten aufweist. Damit können die Parameterstudien, eingeschränkt auf die aktiven Unterräume, das maßgebliche Verhalten der Funktion analysieren, während die Funktion entlang der inaktiven Richtungen entweder vernachlässigt oder mit einfacheren Modellen approximiert wird. In einigen Studien können aktive Unterräume auch zur Bestimmung von Startpunkten für andere Verfahren hilfreich sein, wie beispielsweise bei der Optimierung.

Die Methode der aktiven Unterräume wird zunächst in Kapitel 2 vorgestellt, allerdings werden wir sehen, dass sich das Verfahren nicht ohne weitere Überlegungen in der Praxis

1 Einleitung

anwenden lässt. Aufgrund dessen betrachtet Kapitel 3 die praktische Umsetzung des Ansatzes und behandelt verschiedene Vereinfachungen, mit deren Hilfe sich der Rechenaufwand erheblich reduziert, die allerdings nicht ohne gewisse Einschränkungen in den resultierenden aktiven Unterräumen einhergehen. Zudem werden mögliche Erweiterungen des Vorgehens, beispielsweise auf Funktionen mit mehrdimensionalem Wertebereich, diskutiert. Anschließend stellt Kapitel 4 zwei Anwendungen vor, in denen sich aktive Unterräume berechnen lassen. Des Weiteren werden Hinweise zur Implementierung der präsentierten Algorithmen gegeben und deren Verhalten in den praktischen Beispielen analysiert. Die entsprechende Dimensionsreduktion wird grafisch sichtbar gemacht und die in Abschnitt 3.4 eingeführte Verallgemeinerung zeigt in einer der Anwendungen ihr Potential. Das abschließende Kapitel 5 legt einen Fokus auf die konkrete Verwendung aktiver Unterräume in Optimierungsaufgaben und erläutert einige in diesem Zusammenhang auftretende Schwierigkeiten. Außerdem werden allgemeine Ansätze zum Einsatz in Parameterstudien aufgezeigt. Eine Zusammenfassung der Ergebnisse und Anregungen zu weiteren Forschungsthemen gibt Kapitel 6.

Sämtliche Abbildungen in dieser Arbeit, die nicht entsprechend gekennzeichnet sind, wurden eigenständig konzipiert und mittels \LaTeX umgesetzt.

Eigener Beitrag

Folgende wesentliche Beiträge finden sich in dieser Arbeit:

- Verständnis der theoretischen Hintergründe bei der Berechnung aktiver Unterräume
- Implementierung der in Kapitel 3 vorgestellten Algorithmen
- Entwicklung eines Ansatzes zur Kombination von aktiven Unterräumen und Herleitung einer Optimalitätseigenschaft der berechneten Richtung (siehe Abschnitt 3.4 sowie Abschnitt 4.1.5)
- Anwendung der Algorithmen auf die Beispiele aus Kapitel 4, Einarbeitung in das Programm OPENFOAM[®] [HJF⁺12] sowie Automatisierung der Bestimmung von Auftriebs- und Reibungsbeiwerten bei NACA-Profilen
- Aufarbeitung der Grundlagen zur Nutzung aktiver Unterräume in Parameterstudien und insbesondere der Optimierung
- Anwendung von aktiven Unterräumen auf ein Minimierungsproblem in Zusammenhang zur Strömungssimulation sowie Evaluation der Ergebnisse (siehe Abschnitt 5.2.2)

Danksagung

Zunächst möchte ich mich recht herzlich bei Professor Jochen Garcke für die hervorragende Betreuung, viele wertvolle Anregungen und die Beantwortung sämtlicher Fragen bedanken. Des Weiteren danke ich Professor Michael Griebel für die Übernahme der Zweitkorrektur.

Zudem bedanke ich mich bei Christian Gscheidle für die Unterstützung und erklärende Hinweise beim Umgang mit Strömungssimulationen in OPENFOAM®.

Für das Korrekturlesen der Arbeit beziehungsweise einzelner Beweise danke ich Danny, Jannis, Lars und Max. Für eine wunderbare Zeit während des Studiums bedanke ich mich zudem besonders bei Andrii, Jan, Kai und Niklas.

Ein ganz besonderer Dank gebührt allerdings meiner Familie, insbesondere meinen Eltern und meiner Schwester, für die jahrelange und unermüdliche Unterstützung in jeder Klausurphase, während des Semesters, aber auch in vielen anderen Momenten. Hierfür mein herzlicher Dank. Zudem möchte ich mich bei meiner Freundin Lisa bedanken, die jederzeit motivierende Worte gefunden und mir in vielen Situationen unheimlich geholfen hat.

Kapitel 2

Theoretische Berechnung aktiver Unterräume

Sollen aktive Unterräume in praktischen Anwendungen zur Dimensionsreduktion eingesetzt werden, so müssen diese zunächst berechnet beziehungsweise approximiert und anschließend je nach gewünschter Parameterstudie weiter verwendet werden. Mit dem ersten Schritt, der Ermittlung und Definition der aktiven Richtungen im Parameterraum, beschäftigt sich das folgende Kapitel. Zunächst wird die theoretische Vorgehensweise erläutert, Methoden zur praktischen Umsetzung des Verfahrens werden im nächsten Kapitel diskutiert. Zuallererst soll jedoch geklärt werden, welche Voraussetzungen wir an die betrachtete Funktion stellen und welche Notation verwendet wird.

2.1 Grundlegende Notation und Annahmen

Grundsätzlich betrachten wir eine Funktion

$$f: \mathcal{X} \rightarrow \mathbb{R},$$

wobei $\mathcal{X} \subseteq \mathbb{R}^m$ sei. Wir nehmen an, dass $1 < m < \infty$ gilt. Dabei sollte die Dimension m maximal einige Tausend betragen, um die Berechnung einer Eigenwertzerlegung von einer $m \times m$ -Matrix zu ermöglichen (siehe Kapitel 2.2). Zudem sei f differenzierbar und Lipschitz-stetig und der Gradient von f bezüglich x sei mit $\nabla_x f(x) = \left[\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_m}(x) \right]^T$ bezeichnet. Die genannten Voraussetzungen sind äquivalent zu der Forderung einer beschränkten Norm des Gradienten, das heißt, f ist genau dann differenzierbar und Lipschitz-stetig, wenn

$$\|\nabla_x f(x)\| \leq L \quad \text{für alle } x \in \mathcal{X} \text{ gilt.} \quad (2.1)$$

Falls nicht anders angegeben, handelt es sich bei Vektornormen um die Euklidische Norm und bei Matrixnormen um die Spektralnorm, welche von der euklidischen Norm induziert wird.

Die Menge \mathcal{X} bezeichnet den Parameterraum, die Funktion f ist eine skalare Funkti-

on, die einer gewissen Wahl von Parametern $x \in \mathcal{X}$ einen reellen Wert $f(x)$ zuordnet. Für ein $x \in \mathcal{X}$ schreiben wir $x = [x_1, \dots, x_m]^T$. Typischerweise sind für die einzelnen Parameter x_1, \dots, x_m Werte aus endlichen Intervallen sinnvoll, das heißt, es gibt Vektoren $l \in \mathbb{R}^m$ und $u \in \mathbb{R}^m$ mit $l < u$, die untere beziehungsweise obere Schranken für die Wahl der Parameter enthalten (die Ordnungsrelation zwischen Vektoren ist jeweils komponentenweise zu verstehen, das heißt, es gilt $l < u$ genau dann, wenn $l_i < u_i$ für alle $i = 1, \dots, m$ gilt, wobei $u = [u_1, \dots, u_m]^T$ und $l = [l_1, \dots, l_m]^T$). In diesem Fall ist $\mathcal{X} = \{x \in \mathbb{R}^m : l \leq x \leq u\} =: [l, u]$ ein Quader im Vektorraum \mathbb{R}^m . In einigen Anwendungen kann es allerdings auch sinnvoll sein, \mathcal{X} als Polytop beziehungsweise Polyeder zu wählen, also \mathcal{X} durch ein endliches System linearer Ungleichungen zu beschränken, oder aber $\mathcal{X} = \mathbb{R}^m$ zu setzen.

Um Integrale über \mathcal{X} berechnen zu können, wird \mathcal{X} mit einer Wahrscheinlichkeitsdichte $\rho: \mathcal{X} \rightarrow \mathbb{R}$ versehen. Die nachfolgend betrachteten Integrale sind Lebesgue-Integrale und wir nehmen an, dass ρ eine Wahrscheinlichkeitsdichte bezüglich des Lebesgue-Maßes \mathcal{L} auf \mathbb{R}^m ist. Damit können Integrale auch als Erwartungswerte interpretiert werden. Die tatsächliche Wahl von ρ hängt von der Wahl von \mathcal{X} und der konkreten Anwendung ab. Für den Fall $\mathcal{X} = [l, u]$ mit $l, u \in \mathbb{R}^m$ und $l < u$ wird für ρ häufig eine Gleichverteilung auf dem Quader angenommen, das heißt

$$\rho(x) = \left(\prod_{i=1}^m (u_i - l_i) \right)^{-1} \quad \text{für alle } x \in \mathcal{X},$$

wobei $u = [u_1, \dots, u_m]^T$ und $l = [l_1, \dots, l_m]^T$. Damit ρ wohldefiniert ist, setzen wir $l < u$ voraus, dann ist das Lebesgue-Maß des betrachteten Quaders positiv, das heißt $\mathcal{L}([l, u]) = \prod_{i=1}^m (u_i - l_i) > 0$. Wird keine Gleichverteilung verwendet, so konzentriert sich die Parameterstudie, und entsprechend auch die Berechnung der aktiven Unterräume, auf bestimmte Parameterwerte stärker als auf andere. Im Fall $\mathcal{X} = \mathbb{R}^m$ kann zum Beispiel eine m -dimensionale Normalverteilung zum Einsatz kommen, wobei der Fokus der Berechnungen auf einem gewissen Bereich um den Erwartungswert der Normalverteilung liegt.

Ist aus dem Kontext ersichtlich, über welche Menge integriert wird, so verzichten wir auf die explizite Nennung dieser Menge. Entsprechend verwenden wir häufig $f = f(x)$ und $\rho = \rho(x)$.

Wann immer im Folgenden von Richtungen in einem Vektorraum die Rede ist, wird davon ausgegangen, dass die entsprechenden Vektoren normiert sind.

Zu der Funktion f wollen wir eine Parameterstudie durchführen. Unter Parameterstudien verstehen wir beispielsweise Optimierung, Integration, Inversion, Regression oder ähnliche Studien, die in einer gewissen Weise die Zusammenhänge zwischen gewählten Parameterwerten und der skalaren Zielfunktion f analysieren. Beispielsweise kann mit Hilfe von Integralen der Durchschnitt der Funktionswerte von f über \mathcal{X} beschrieben werden. Optimierung kann andererseits helfen, den Wertebereich von f anzugeben, indem minimaler und maximaler Funktionswert bestimmt werden.

Der folgende Abschnitt stellt die Idee der aktiven Unterräume vor. Insbesondere wird erläutert, worum es sich bei aktiven Unterräumen handelt und wie diese in der Theorie

bestimmt werden.

2.2 Theoretischer Ansatz

Die Grundlage zur Berechnung aktiver Unterräume stammt von Russi [Rus10] und beruht auf der Idee, die quadrierte Richtungsableitung von f in eine feste Richtung w zu betrachten und diese über den gesamten Parameterraum \mathcal{X} zu mitteln. Es wird also das Integral, beziehungsweise der Erwartungswert, von quadrierten Richtungsableitungen untersucht, um festzustellen, entlang welcher Richtungen sich der Funktionswert von f , im Vergleich zu den orthogonalen Richtungen, besonders stark verändert. Russi entwickelte eine Methode, um ein orthonormales System von Richtungen zu berechnen, das die Bestimmung von aktiven Unterräumen ermöglicht. Dieses Vorgehen wird im Folgenden eingeführt und näher erläutert. Die Darstellungen in diesem Abschnitt orientieren sich an Kapitel 3.2 aus [Con15].

2.2.1 Unzentrierte Kovarianzmatrix der ersten partiellen Ableitungen

Es sei $C \in \mathbb{R}^{m \times m}$ definiert als

$$C := \int_{\mathcal{X}} \left[(\nabla_x f(x)) (\nabla_x f(x))^T \right] \rho(x) dx. \quad (2.2)$$

Die Einträge von $C = [c_{ij}]_{i,j=1}^m$ sind also für alle $i, j = 1, \dots, m$ gegeben durch

$$c_{ij} = \int_{\mathcal{X}} \frac{\partial f(x)}{\partial x_i} \frac{\partial f(x)}{\partial x_j} \rho(x) dx. \quad (2.3)$$

Die Matrix C entspricht dem über den gesamten Parameterraum integrierten und mit der Wahrscheinlichkeitsdichte ρ gewichteten dyadischen Produkt des Gradienten von f mit sich selbst. Sie ist wohldefiniert, denn mit der Voraussetzung aus (2.1), dass die Norm des Gradienten von f durch L beschränkt ist, gilt für alle $i, j = 1, \dots, m$ die Ungleichung

$$\begin{aligned} |c_{ij}| &= \left| \int_{\mathcal{X}} \frac{\partial f(x)}{\partial x_i} \frac{\partial f(x)}{\partial x_j} \rho(x) dx \right| \\ &\leq \int_{\mathcal{X}} \left| \frac{\partial f(x)}{\partial x_i} \right| \left| \frac{\partial f(x)}{\partial x_j} \right| \rho(x) dx \\ &\leq \int_{\mathcal{X}} L^2 \rho(x) dx \\ &= L^2 \int_{\mathcal{X}} \rho(x) dx \\ &= L^2. \end{aligned}$$

Bei der Abschätzung wurde ausgenutzt, dass ρ eine Wahrscheinlichkeitsdichte ist und dass sich der maximale Eintrag des Gradienten ebenfalls durch die Euklidische Norm des Gra-

dienten abschätzen lässt, das heißt es gilt

$$\max_{i=1,\dots,m} \left| \frac{\partial f(x)}{\partial x_i} \right| = \|\nabla_x f(x)\|_\infty \leq \|\nabla_x f(x)\|_2 \leq L \quad \text{für alle } x \in \mathcal{X}.$$

An dieser Stelle würde es bereits genügen, wenn die Norm des Gradienten fast-sicher (bezüglich des mit der Dichte ρ gewichteten Lebesgue-Maßes) beschränkt wäre. Solche Überlegungen könnten allerdings bereits in den Definitionsbereich \mathcal{X} von f mit einfließen.

Eine alternative Interpretation der Integrale als Erwartungswerte wurde bereits in Abschnitt 2.1 angedeutet. Diese soll nun präzisiert werden. Dazu sei S eine Zufallsvariable mit Werten aus \mathbb{R}^m , deren Verteilung die Dichte ρ besitzt. Dann gilt

$$C = \mathbb{E} \left[(\nabla_x f(S)) (\nabla_x f(S))^T \right].$$

Damit kann C als unzentrierte Kovarianzmatrix der ersten partiellen Ableitungen gedeutet werden [CDW14], siehe auch Formel (2.3). Diese Definition der Matrix C wurde von Samarov im Bereich der parameterfreien Regression entwickelt [Sam93].

Wir fassen nun einige Eigenschaften von C in folgendem Lemma zusammen:

Lemma 2.1. *Die Matrix $C \in \mathbb{R}^{m \times m}$, definiert als*

$$C := \int_{\mathcal{X}} \left[(\nabla_x f(x)) (\nabla_x f(x))^T \right] \rho(x) dx,$$

ist symmetrisch und positiv-semidefinit.

Beweis. Die Matrix C ist symmetrisch, denn mit (2.3) folgt

$$c_{ij} = \int_{\mathcal{X}} \frac{\partial f(x)}{\partial x_i} \frac{\partial f(x)}{\partial x_j} \rho(x) dx = \int_{\mathcal{X}} \frac{\partial f(x)}{\partial x_j} \frac{\partial f(x)}{\partial x_i} \rho(x) dx = c_{ji}$$

für alle $i, j = 1, \dots, m$. Zudem ist C positiv-semidefinit, da für alle $w \in \mathbb{R}^m$ gilt

$$w^T C w = \int_{\mathcal{X}} \left[w^T (\nabla_x f(x)) (\nabla_x f(x))^T w \right] \rho(x) dx = \int_{\mathcal{X}} \left(w^T (\nabla_x f(x)) \right)^2 \rho(x) dx \geq 0.$$

Bei der Abschätzung wurde ausgenutzt, dass ρ eine Wahrscheinlichkeitsdichte und damit nicht-negativ ist. ■

Da C nach Lemma 2.1 symmetrisch und positiv-semidefinit ist, besitzt C nur reelle, nicht-negative Eigenwerte. Zudem existiert eine orthonormale Basis aus Eigenvektoren, das heißt, wir können eine Eigenwertzerlegung von C mit der Form

$$C = W \Lambda W^T$$

berechnen. Hierbei ist $W = [w_1, \dots, w_m] \in \mathbb{R}^{m \times m}$ eine orthogonale Matrix, deren Spalten $w_1, \dots, w_m \in \mathbb{R}^m$ den Eigenvektoren von C entsprechen, und $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_m) \in \mathbb{R}^{m \times m}$ die Diagonalmatrix mit den entsprechenden absteigend sortierten Eigenwerten von C . Für die Eigenwerte von C gilt also $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m \geq 0$. Die Eigenwertzerlegung

kann in diesem Fall auch mit Hilfe der Singulärwertzerlegung berechnet werden (siehe Abschnitt 3.3).

Das folgende Lemma von Constantine [Con15] formuliert den Zusammenhang zwischen Richtungsableitungen von f und den Eigenpaaren von C :

Lemma 2.2 ([Con15], Lemma 3.1). *Die mittlere quadratische Richtungsableitung von f bezüglich des Eigenvektors w_i entspricht dem zugehörigen Eigenwert λ_i , das heißt*

$$\int ((\nabla_x f)^T w_i)^2 \rho \, dx = \lambda_i \quad \text{für alle } i = 1, \dots, m.$$

Beweis. Die Eigenvektoren w_1, \dots, w_m von C sind normiert, das heißt es gilt $w_i^T w_i = 1$ für alle $i = 1, \dots, m$. Nun folgt mit der Definition von C und dem Rayleigh-Quotienten, dass

$$\lambda_i = \frac{w_i^T C w_i}{w_i^T w_i} = w_i^T C w_i = w_i^T \left(\int (\nabla_x f)(\nabla_x f)^T \rho \, dx \right) w_i = \int ((\nabla_x f)^T w_i)^2 \rho \, dx$$

für alle $i = 1, \dots, m$ gilt. ■

Bemerkung 2.1. Der Wert $(\nabla_x f(x))^T w_i$ entspricht der Richtungsableitung der Funktion f an der Stelle $x \in \mathcal{X}$ in Richtung w_i . Da die Eigenvektoren aus der Singulärwertzerlegung allesamt normiert sind, lassen sich nach Lemma 2.2 anstelle der gemittelten Richtungsableitungen in Richtung w_1, \dots, w_m auch die Eigenwerte $\lambda_1, \dots, \lambda_m$ vergleichen. ◁

2.2.2 Bestimmung aktiver und inaktiver Richtungen

Im letzten Abschnitt haben wir gesehen, dass eine orthonormale Basis des Vektorraums \mathbb{R}^m , bestehend aus Eigenvektoren der Matrix C , existiert. Zudem trifft Lemma 2.2 eine Aussage über den Zusammenhang der Eigenwerte von C und Richtungsableitungen von f in Richtung von Eigenvektoren. Diese beiden Resultate motivieren den Ansatz zur Berechnung von aktiven und inaktiven Richtungen im Parameterraum, der in diesem Abschnitt diskutiert wird.

Zunächst betrachten wir noch einmal Lemma 2.2 für den Fall, dass es ein $n \in \mathbb{N}$ mit $1 \leq n < m$ gibt, sodass $\lambda_{n+1} = \lambda_{n+2} = \dots = \lambda_m = 0$ gilt. Die kleinsten $m - n$ Eigenwerte seien also bereits gleich null. Dann gilt jedoch auch

$$(\nabla_x f)^T w_{n+1} = (\nabla_x f)^T w_{n+2} = \dots = (\nabla_x f)^T w_m = 0$$

fast-sicher, das heißt die Richtungsableitungen von f in Richtung der Eigenvektoren w_{n+1}, \dots, w_m verschwinden außerhalb einer Nullmenge. Da die Funktion f differenzierbar ist, ist f damit entlang der Richtungen w_{n+1}, \dots, w_m konstant. Es folgt für alle $x \in \mathcal{X}$ und $\tau_{n+1}, \dots, \tau_m \in \mathbb{R}$ mit $x + \tau_{n+1} w_{n+1} + \dots + \tau_m w_m \in \mathcal{X}$, dass gilt

$$f(x) = f(x + \tau_{n+1} w_{n+1} + \dots + \tau_m w_m).$$

Dieses Ergebnis kann nun zur Dimensionsreduktion genutzt werden. Die Funktionswerte von f sind unabhängig von den Richtungen w_{n+1}, \dots, w_m , das heißt, sämtliche Parameterstudien von f können sich auf die zu w_{n+1}, \dots, w_m orthogonalen Richtungen beschränken,

also auf w_1, \dots, w_n . Durch eine geeignete Koordinatentransformation lässt sich somit sogar die Dimension des Parameterraumes auf $n < m$ reduzieren. Diese Transformation ist durch die Eigenvektormatrix W aus der Eigenwertzerlegung von C gegeben. Die Matrix W entspricht einer Rotation (beziehungsweise einer Drehspiegelung, falls $\det(W) = -1$ gilt), welche die Einheitsvektoren auf die entsprechenden Eigenvektoren abbildet. Die transformierte Basis besteht damit aus den Vektoren w_1, \dots, w_m . Die Vektoren w_1, \dots, w_n bestimmen dabei den *aktiven Unterraum*, die Vektoren w_{n+1}, \dots, w_m bilden inaktive Richtungen. Der aktive Unterraum ist insbesondere also keine Teilmenge des Parameterraumes, sondern eine *Menge von Richtungen* im Parameterraum.

In den meisten Anwendungen kann nicht davon ausgegangen werden, dass Eigenwerte von C bereits verschwinden. Genauso wenig lässt sich annehmen, dass die Eigenwerte sehr klein sind. Trotzdem kann ein aktiver Unterraum existieren, der das Verhalten der Funktion f hinreichend gut beschreibt, um Parameterstudien in geeigneter Weise auf die Richtungen des aktiven Unterraums zu beschränken. Hierzu werden aktive von inaktiven Richtungen unterschieden, indem die Lücken zwischen den Eigenwerten von C betrachtet werden. Es wird nach der größten Lücke zweier, in der absteigenden Sortierung, benachbarter Eigenwerte gesucht (siehe Korollar 3.7). Die Diagonalmatrix Λ und entsprechend die Eigenvektormatrix W werden also in Blöcke eingeteilt,

$$\Lambda = \begin{bmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{bmatrix}, \quad W = \begin{bmatrix} W_1 & W_2 \end{bmatrix}.$$

Die Matrix $\Lambda_1 = \text{diag}(\lambda_1, \dots, \lambda_n) \in \mathbb{R}^{n \times n}$ enthält die n größten Eigenwerte. Dementsprechend besteht $W_1 \in \mathbb{R}^{m \times n}$ aus den zugehörigen n Eigenvektoren. Die Dimension n des aktiven Unterraums wird, wie oben beschrieben, durch den größten Abstand aufeinander folgender Eigenwerte bestimmt. Es wird also $1 \leq n < m$ so gewählt, dass

$$\lambda_n - \lambda_{n+1} \geq \lambda_i - \lambda_{i+1} \geq 0 \quad (2.4)$$

für alle $i = 1, \dots, m-1$ gilt. Damit können wir n auch schreiben als

$$n = \underset{k=1, \dots, m-1}{\operatorname{argmax}} \lambda_k - \lambda_{k+1}. \quad (2.5)$$

Ist dieses Maximum nicht eindeutig, so wird n minimal mit der Eigenschaft (2.4) gewählt.

Wir können nun jedes $x \in \mathcal{X}$ in der Basis darstellen, die aus den Spalten von W besteht. Zudem können wir die Basisvektoren in einen aktiven und inaktiven Anteil aufteilen. Das Ziel ist es $y \in \mathbb{R}^n$ und $z \in \mathbb{R}^{m-n}$ zu finden, sodass

$$x = W_1 y + W_2 z$$

gilt. Hierzu setzen wir $y = W_1^T x$ als *aktive Variablen* und entsprechend $z = W_2^T x$ als *inaktive Variablen*. Dann gilt

$$x = WW^T x = \begin{bmatrix} W_1 & W_2 \end{bmatrix} \begin{bmatrix} W_1^T \\ W_2^T \end{bmatrix} x = W_1 W_1^T x + W_2 W_2^T x = W_1 y + W_2 z.$$

In obiger Rechnung haben wir ausgenutzt, dass W orthogonal ist und damit $WW^T = I_m$ gilt, wobei $I_m \in \mathbb{R}^{m \times m}$ die Einheitsmatrix bezeichnet.

Bemerkung 2.2. An dieser Stelle sei angemerkt, dass $W_1 W_1^T$ einen Orthogonalprojektor bezüglich des Standardskalarproduktes auf den von den Spalten der Matrix W_1 aufgespannten linearen Unterraum $\text{im}(W_1)$ darstellt. Dies folgt daraus, dass W_1 eine spaltenorthogonale Matrix ist, das heißt $W_1^T W_1 = I_n$ gilt. Damit ergeben sich die Beziehungen

$$(W_1 W_1^T)^2 = W_1 W_1^T W_1 W_1^T = W_1 W_1^T \quad (2.6)$$

$$\text{und } (W_1 W_1^T)^T = (W_1^T)^T W_1^T = W_1 W_1^T, \quad (2.7)$$

welche gerade den Bedingungen für einen Orthogonalprojektor entsprechen. Zudem gilt $\text{im}(W_1 W_1^T) = \text{im}(W_1)$. Somit stimmt das Bild des Projektors $W_1 W_1^T$ mit dem Bild von W_1 überein. Diese Eigenschaft wird in Abschnitt 3.1.2 und Abschnitt 5.1 von Bedeutung sein. \triangleleft

2.2.3 Beispiel

Wir betrachten nun ein einfaches Beispiel zur Berechnung aktiver Unterräume:

Es sei $f: \mathcal{X} \rightarrow \mathbb{R}$ mit $\mathcal{X} = [-1, 1]^2$ gegeben als

$$f(x) = \cos(a^T x),$$

wobei $a \in \mathbb{R}^2 \setminus \{0\}$ beliebig sei. Die folgende Abbildung zeigt den Graphen von f für die Wahl $a = [2 \ 5]^T$:

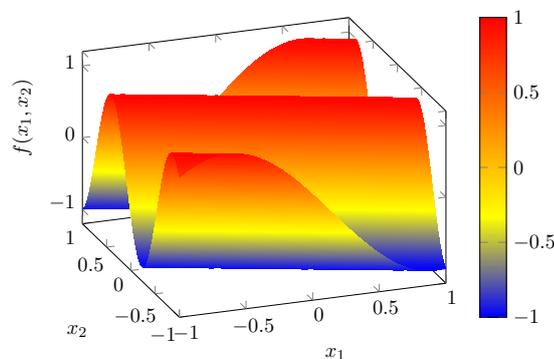


Abbildung 2.1: Dreidimensionale Abbildung der Funktion $f(x_1, x_2) = \cos(2x_1 + 5x_2)$

Für den Gradienten von f an der Stelle $x = [x_1 \ x_2]^T \in \mathcal{X}$ gilt

$$\nabla_x f(x) = a \cdot \cos(a^T x).$$

Als Wahrscheinlichkeitsdichte nehmen wir eine Gleichverteilung auf \mathcal{X} an, wir setzen also

$\rho(x) = \frac{1}{4}$ für alle $x \in \mathcal{X}$. Damit ergibt sich die Matrix C als

$$\begin{aligned} C &= \int_{\mathcal{X}} [(\nabla_x f(x))(\nabla_x f(x))^T] \rho(x) dx \\ &= \int_{\mathcal{X}} \frac{1}{4} [(a \cdot \cos(a^T x))(a \cdot \cos(a^T x))^T] dx \\ &= \frac{1}{4} \int_{\mathcal{X}} [\cos(a^T x)^2 \cdot aa^T] dx \\ &= \left(\frac{1}{4} \int_{\mathcal{X}} \cos(a^T x)^2 dx \right) \cdot aa^T \\ &= K \cdot aa^T, \end{aligned}$$

mit $K = \frac{1}{4} \int_{\mathcal{X}} \cos(a^T x)^2 dx$.

Es sei nun $b \in \mathbb{R}^2$ mit $a^T b = 0$, das heißt, b ist orthogonal zu a bezüglich des Standardskalarproduktes. Dann gelten die Beziehungen

$$\begin{aligned} C \cdot a &= K \cdot aa^T a = (K \cdot a^T a) \cdot a \\ \text{und } C \cdot b &= K \cdot aa^T b = 0. \end{aligned}$$

Damit ist a ein Eigenvektor der Matrix C zum Eigenwert $K \cdot a^T a$ und b ist ein Eigenvektor von C zum Eigenwert 0. Aus der Diskussion in Abschnitt 2.2.2 für den Fall, dass 0 ein Eigenwert von C ist, folgt nun, dass die Funktion f in Richtung b konstant ist. Dieses Verhalten illustrieren folgende Grafiken, bei denen wie in Abbildung 2.1 der Vektor $a = [2 \ 5]^T$ gewählt wurde. Die normierten Eigenvektoren von C sind in diesem Fall $w_1 = \frac{1}{\sqrt{29}} [2 \ 5]^T$, sowie $w_2 = \frac{1}{\sqrt{29}} [-5 \ 2]^T$.

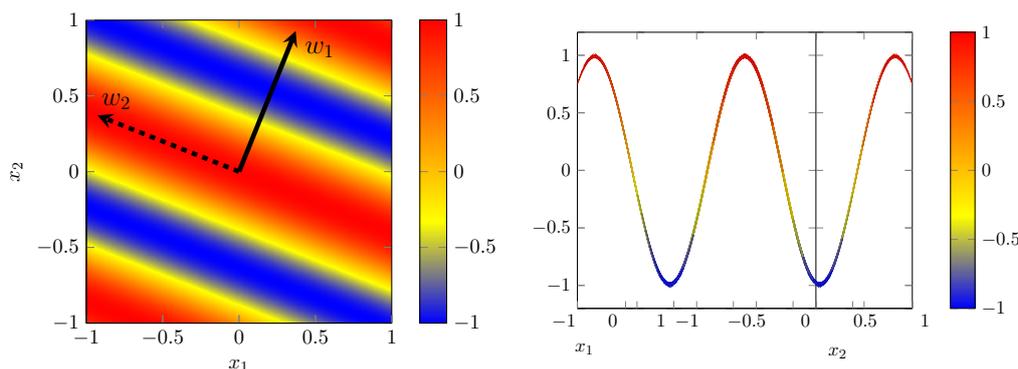


Abbildung 2.2: Die Funktion $f(x_1, x_2) = \cos(2x_1 + 5x_2)$ von oben (links) und von der Seite (rechts)

Die obigen Abbildungen demonstrieren, wie sich die Dimension des Parameterraums in diesem Beispiel reduzieren lässt. Die linke Grafik zeigt die Funktion f , wobei von oben auf die x_1 - x_2 -Ebene geblickt wird. Anhand der Farben lässt sich erkennen, dass die Funktionswerte von f entlang der Richtung w_2 konstant sind und sich nur entlang von w_1 verändern. Wird die durch die Eigenvektoren der Matrix C definierte Rotation auf f und den Parameter-

raum \mathcal{X} angewandt, so lässt sich eine eindimensionale Struktur ausmachen (rechte Grafik; die Rotation ist insbesondere durch die Achsenbeschriftung erkennbar). Mit Hilfe einer geeigneten Transformation lässt sich die Funktion f somit in eine Funktion überführen, die nur von einer Variablen abhängt.

Bemerkung 2.3. Das vorangehende Beispiel kann noch weiter verallgemeinert werden (siehe [Sam93], Abschnitt 2, Modell e und f). Hierzu setzen wir

$$f(x) = g_1(a_1^T x) + g_2(a_2^T x) + \cdots + g_k(a_k^T x)$$

mit $g_i: \mathcal{X} \rightarrow \mathbb{R}$ differenzierbar und $a_i \in \mathbb{R}^m$ für $i = 1, \dots, k$. Auch in diesem Fall ergibt sich eine recht einfache Struktur der Matrix C , die wieder als Linearkombination der äußeren Produkte der Vektoren a_i gegeben ist. Samarov betrachtet die Matrix C , sowie einige weitere Funktionen, die aus (höheren) Ableitungen von f entstehen. \triangleleft

2.2.4 Zusammenfassung

In diesem Kapitel haben wir die grundlegende Idee sowie das theoretische Vorgehen bei der Berechnung aktiver Unterräume kennengelernt. Der folgende Algorithmus fasst den Ansatz noch einmal zusammen:

Algorithmus 1 Aktive Unterräume (theoretisch)

- 1: Berechne die Matrix $C = \int [(\nabla_x f)(\nabla_x f)^T] \rho dx$
 - 2: Ermittle die Eigenwertzerlegung $C = W\Lambda W^T$, wobei W die Matrix der orthonormalen Eigenvektoren bezeichnet und $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_m)$, die Diagonalmatrix mit den entsprechenden Eigenwerten darstellt ($\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m \geq 0$)
 - 3: Suche nach dem größten Abstand zweier aufeinander folgender Eigenwerte
 - 4: Bestimme dementsprechend die Dimension des aktiven Unterraums n , die Matrix der aktiven Richtungen W_1 und die Matrix der inaktiven Richtungen W_2
-

Zur praktischen Umsetzung ist dieser Algorithmus allerdings nicht geeignet, denn es müssen Integrale über einem hochdimensionalen Parameterraum berechnet werden. Zudem wird hier vorausgesetzt, dass Gradienten der Funktion f vorliegen. Um in der Praxis trotzdem aktive Unterräume nutzen zu können, betrachten wir im folgenden Kapitel Ansätze, wie mit den soeben beschriebenen Schwierigkeiten umgegangen werden kann. Wir werden unterschiedliche Algorithmen kennenlernen und einige Resultate zum Verhalten dieser Algorithmen begutachten.

Umsetzung des Verfahrens

Die theoretische Grundlage zur Berechnung aktiver Unterräume aus Abschnitt 2.2 lässt sich nicht ohne Weiteres auf praktische Aufgabenstellungen übertragen. Es sind zunächst einige Fragen zu klären:

1. Wie kann die Matrix C berechnet beziehungsweise approximiert werden?
2. Muss der Gradient der Funktion f bekannt sein und für alle $x \in \mathcal{X}$ berechnet werden können?
3. Wie lässt sich die Eigenwertzerlegung von C bestimmen?

All diese Fragen sollen in den folgenden Abschnitten geklärt werden. Weitere Hinweise zur Anwendung in praktischen Problemen finden sich auch in Kapitel 4, das Beispiele zur Bestimmung aktiver Unterräume enthält. In Abschnitt 3.4 wird zudem eine Möglichkeit erläutert, wie sich aktive Unterräume auf mehrere Funktionen verallgemeinern lassen.

3.1 Approximation der Matrix C

In der Praxis ist die Funktion f das Ergebnis einer aufwändigen Simulation oder auf andere Art und Weise aus aufwändigen Rechnungen entstanden. Insbesondere ist es häufig nur möglich, den Wert $f(x)$ für ein gegebenes $x \in \mathcal{X}$ zu berechnen, nicht aber eine explizite Formel für die Funktion f anzugeben, genauso wenig wie für ihren Gradienten. Sollen aktive Unterräume in der Praxis genutzt werden, so muss allerdings die Matrix C berechnet werden. Diese Matrix ist das Ergebnis eines Integrals über einem hochdimensionalen Parameterraum. Standardverfahren für eindimensionale numerische Integration, beispielsweise Quadraturformeln mit deterministischen Wahlen von Stützstellen, leiden unter dem *Fluch der Dimension*. Daher sind sie in hochdimensionalen Räumen unpraktisch. Insbesondere gehen wir davon aus, dass eine einzige Funktionsauswertung (beziehungsweise die Berechnung des Gradienten an einem einzigen Punkt) bereits sehr aufwändig ist. Dies erschwert die numerische Integration zusätzlich. Um Integrale mit hochdimensionalem Parameterraum trotzdem approximieren zu können, finden häufig Monte-Carlo-Verfahren Verwendung. Die (fast-sichere) Konvergenz dieser Verfahren beruht auf dem starken Ge-

setz der großen Zahlen. An dieser Stelle tritt also wieder die Deutung der Matrix C als Erwartungswert auf.

Wir nehmen nun zunächst an, dass uns ein Black-Box-Algorithmus zur Verfügung steht, der den Gradienten von f an einer Stelle $x \in \mathcal{X}$ zurückgibt. Zudem gehen wir davon aus, dass wir die Eigenwertzerlegung einer $m \times m$ -Matrix exakt berechnen können. Der Algorithmus approximiert zunächst die Matrix C durch eine endliche Summe dyadischer Produkte von Samples des Gradienten mit sich selbst. Die Approximation sei mit \hat{C} bezeichnet. Anschließend werden die aktiven Unterräume wie in Algorithmus 1 in den Schritten 2-4 bestimmt, allerdings unter Verwendung der Matrix \hat{C} . Es ergibt sich folgender Algorithmus zur Ermittlung aktiver Unterräume:

Algorithmus 2 Aktive Unterräume (praktisch, mit Gradienten)

- 1: Ziehe M unabhängige Samples $\{x_i\}_{i=1}^M$ bezüglich der Dichte ρ
 - 2: Nutze die Black-Box um $\nabla_i := \nabla_x f(x_i)$ für $i = 1, \dots, M$ zu berechnen
 - 3: Approximiere C durch $\hat{C} := \frac{1}{M} \sum_{i=1}^M \nabla_i \nabla_i^T$
 - 4: Ermittle die Eigenwertzerlegung $\hat{C} = \hat{W} \hat{\Lambda} \hat{W}^T$, wobei \hat{W} die Matrix der orthonormalen Eigenvektoren bezeichnet und $\hat{\Lambda} = \text{diag}(\hat{\lambda}_1, \dots, \hat{\lambda}_m)$, die Diagonalmatrix mit den entsprechenden Eigenwerten darstellt ($\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \dots \geq \hat{\lambda}_m \geq 0$)
 - 5: Suche nach dem größten Abstand zweier aufeinander folgender Eigenwerte
 - 6: Bestimme dementsprechend die Dimension des aktiven Unterraums n , die Matrix der aktiven Richtungen \hat{W}_1 und die Matrix der inaktiven Richtungen \hat{W}_2
-

Im Allgemeinen unterscheiden sich die von Algorithmus 1 und Algorithmus 2 berechneten Eigenwerte und Eigenvektoren voneinander, also sind auch die aktiven Unterräume nicht identisch. Daher stellt sich die Frage, ob die durch Algorithmus 2 bestimmten aktiven Unterräume überhaupt zur Verwendung in der Praxis geeignet sind. Insbesondere ist für uns von Interesse, wie stark die Eigenwerte von C und \hat{C} sowie die aktiven Unterräume, definiert durch W_1 und \hat{W}_1 , voneinander abweichen. Es lässt sich zeigen, dass wir durch eine hinreichende Wahl der Anzahl M an zu berechnenden Samples sowohl den Abstand der Eigenwerte als auch der aktiven Unterräume kontrollieren können. Die entsprechenden Aussagen von Constantine und Gleich (insbesondere Korollar 3.3 aus [CG15] und Korollar 3.10 aus [Con15]) sollen nun vorgestellt werden. Es wird sich allerdings herausstellen, dass die geeignete Wahl von M von Eigenschaften der Funktion f abhängt, die in der Praxis meist nicht bekannt sind. Kapitel 4 wird zeigen, dass sich trotz alledem zweckmäßige Ergebnisse in praktischen Anwendungen erzielen lassen.

Bemerkung 3.1. Quasi-Monte-Carlo-Verfahren nutzen, im Gegensatz zu den pseudozufälligen Punkten in Monte-Carlo-Verfahren, deterministische Sequenzen von Punkten im Parameterraum, beispielsweise die Halton-Sequenz oder die Sobol-Sequenz. Diese deterministischen Punkte werden so gewählt, dass der Parameterraum möglichst gleichmäßig abgedeckt wird. Daher können Quasi-Monte-Carlo-Methoden in unserem Fall nur dann ohne erheblichen Mehraufwand eingesetzt werden, wenn der Parameterraum \mathcal{X} ein endlicher Quader und ρ die Gleichverteilung auf diesem Quader ist. Der Parameterraum $\mathcal{X} = [l, u]$ mit $l < u$ und $l, u \in \mathbb{R}^m$ kann in diesem Fall durch geeignete Skalierung und Verschiebung

bijektiv in den Einheitswürfel $[0, 1]^m$ abgebildet werden. Auf diesem sind die oben genannten Sequenzen definiert. Eventuell führen Quasi-Monte-Carlo-Verfahren zu einem besseren Konvergenz-Verhalten der Approximation der Integrale als die Monte-Carlo-Methode aus Algorithmus 2. Für weitere Details zu Quasi-Monte-Carlo-Methoden siehe beispielsweise [Nie92] oder [Caf98]. \triangleleft

3.1.1 Abstand der Eigenwerte

Wir beginnen damit, eine Abschätzung für den Abstand der Eigenwerte von C und \hat{C} anzugeben. Der Beweis des folgenden Satzes nutzt eine Matrix-Bernstein-Ungleichung die sich in [GT11] findet.

Satz 3.1 ([CG15], Theorem 3.1). *Es sei $\|\nabla_x f(x)\| \leq L$ für alle $x \in \mathcal{X}$. Dann gilt für alle $0 < \varepsilon \leq 1$, dass*

$$\mathbb{P} \left[\hat{\lambda}_k \geq (1 + \varepsilon)\lambda_k \right] \leq (m - k + 1) \exp \left(\frac{-M\lambda_k \varepsilon^2}{4L^2} \right)$$

und

$$\mathbb{P} \left[\hat{\lambda}_k \leq (1 - \varepsilon)\lambda_k \right] \leq k \exp \left(\frac{-M\lambda_k^2 \varepsilon^2}{4\lambda_1 L^2} \right).$$

Bemerkung 3.2. Der exponentielle Abfall der Wahrscheinlichkeiten in M entspricht dem grundlegenden Verhalten von Monte-Carlo-Verfahren, deren Konvergenz auf starken Gesetzen großer Zahlen beruht. Abschätzungen von Wahrscheinlichkeiten großer Abweichungen vom Gesetz der großen Zahlen im Fall reellwertiger Zufallsvariablen liefert beispielsweise der Satz von Chernoff [SW95]. Auch hier ist der Abfall der Wahrscheinlichkeiten exponentiell in der Anzahl der Samples. Die hier betrachteten Zufallsvariablen sind tatsächlich die Eigenwerte der zufälligen Matrix \hat{C} aus Algorithmus 2. \triangleleft

Dieser Satz liefert zwei Schranken für die Wahrscheinlichkeit, dass sich der k -te Eigenwert von \hat{C} um mehr als ε vom k -ten Eigenwert von C unterscheidet, wobei sowohl eine Abweichung nach unten als auch nach oben betrachtet wird. Je mehr Samples zur Berechnung der Approximation \hat{C} genutzt werden, desto geringer ist die Wahrscheinlichkeit, dass die Eigenwerte von \hat{C} stark von den Eigenwerten von C abweichen. Insbesondere besagt der Satz, dass die Wahrscheinlichkeit einer großen Abweichung vom gesuchten Wert λ_k exponentiell in M abfällt. Die beiden Schranken kombiniert folgendes Korollar. Es gibt an, wie die Anzahl M der Samples gewählt werden sollte, um eine relative Genauigkeit von ε im k -ten Eigenwert von \hat{C} zu erreichen.

Korollar 3.2 ([CG15], Korollar 3.3). *Wird M als*

$$M = \Omega \left(\frac{L^2 \lambda_1}{\lambda_k^2 \varepsilon^2} \log(m) \right)$$

gewählt, so gilt für alle $\varepsilon \in (0, 1]$ mit hoher Wahrscheinlichkeit

$$\left| \hat{\lambda}_k - \lambda_k \right| \leq \varepsilon \lambda_k.$$

Bemerkung 3.3. Die Formulierung „mit hoher Wahrscheinlichkeit“ meint hier, dass für M gegen Unendlich

$$\mathbb{P} \left[\left| \hat{\lambda}_k - \lambda_k \right| \leq \varepsilon \lambda_k \right] \longrightarrow 1$$

gilt, siehe auch den Beweis von Korollar 3.5. Tatsächlich ist die Konvergenz sogar exponentiell in M , vergleiche Satz 3.1. \triangleleft

In den meisten Anwendungen sind die Lipschitz-Konstante L sowie die Eigenwerte $\lambda_1, \dots, \lambda_m$ der Matrix C nicht bekannt. Daher kann die Wahl von M aus Korollar 3.2 nicht unmittelbar in der Praxis genutzt werden. Am Ende von Abschnitt 3.1.2 werden wir sehen, wie M in praktischen Anwendungen gewählt werden sollte, um adäquate Resultate zu erhalten.

3.1.2 Abstand der Unterräume

Der vorangehende Abschnitt hat gezeigt, dass sich die Eigenwerte von C und \hat{C} wie gewünscht verhalten. Wird die Anzahl M der genutzten Samples erhöht, nähern sich die Eigenwerte der beiden Matrizen einander an. Für die Methode der aktiven Unterräume sind allerdings insbesondere die Eigenvektoren der beiden Matrizen von Interesse, diese definieren nämlich aktive sowie inaktive Richtungen. Daher muss neben dem Abstand der Eigenwerte auch der Abstand der durch C beziehungsweise \hat{C} erzeugten aktiven Unterräume untersucht werden. Wir werden uns nun ansehen, wie der Abstand zwischen zwei Unterräumen definiert ist und wie diese Definition anschaulich zu verstehen ist. An dieser Stelle betrachten wir nur den reellen Fall, der für unsere Anwendung von Bedeutung ist. Stewart beschreibt in [Ste73] allgemeinere Versionen der folgenden Aussagen.

Definition 3.1 (vergleiche [Ste73], Definition 2.1). *Seien $U, V \subseteq \mathbb{R}^m$ Untervektorräume des Vektorraums \mathbb{R}^m , wobei $U \neq \{0\}$ und $V \neq \{0\}$ seien. Der Abstand $\text{dist}(U, V)$ zwischen U und V ist definiert als*

$$\text{dist}(U, V) := \max \left\{ \sup_{\substack{\|u\|=1 \\ u \in U}} \inf_{v \in V} \|u - v\|, \sup_{\substack{\|v\|=1 \\ v \in V}} \inf_{u \in U} \|v - u\| \right\}. \quad (3.1)$$

Bemerkung 3.4. In der Definition ist es wichtig zu fordern, dass beide Untervektorräume nicht dem Null-Vektorraum entsprechen. In diesem Fall sollte nämlich $\text{dist}(U, \{0\}) = 0$ für jeden Unterraum $U \subseteq \mathbb{R}^m$ gesetzt werden. Zudem ist die oben definierte Distanzfunktion im allgemeinen keine Metrik, die Dreiecksungleichung ist nicht zwangsläufig erfüllt. In dem für uns interessanten Fall, dass die verwendete Norm die Euklidische Norm ist, lässt sich zeigen, dass (3.1) eine Metrik definiert [Ste73]. \triangleleft

Der folgende Satz stammt in dieser Formulierung von Stewart [Ste73]. Grundlage für den Satz sind allerdings Aussagen von Kato [Kat95]. Dieser verwendet eine äquivalente Definition des Abstandsbegriffs von Unterräumen und beschreibt detailliert die Verwandlung der Distanzfunktion in eine Metrik. Für uns ist allerdings vielmehr der Zusammenhang zwischen dem Distanzbegriff und Orthogonalprojektoren auf die Unterräume von Bedeutung:

Satz 3.3 ([Ste73], Theorem 2.2). Seien \mathfrak{P}_U und \mathfrak{P}_V Orthogonalprojektoren bezüglich des Standardskalarprodukts auf die Unterräume U und V des Vektorraums \mathbb{R}^m . Wird in Definition 3.1 die Euklidische Norm zur Definition des Abstandes genutzt, dann gilt

$$\text{dist}(U, V) = \|\mathfrak{P}_U - \mathfrak{P}_V\|_2 = \sup_{\substack{\|x\|_2=1 \\ x \in \mathbb{R}^m}} \|\mathfrak{P}_U(x) - \mathfrak{P}_V(x)\|_2.$$

Bemerkung 3.5. Die Definition des Abstandes zweier Untervektorräume $U, V \subset \mathbb{R}^2$ wird durch die folgende Zeichnung illustriert. Hierbei sind x_1 und x_2 sowie x_3 und x_4 jeweils normierte Vektoren aus U und V , das heißt $x_1, x_2 \in U \cap \mathcal{S}^1$ und $x_3, x_4 \in V \cap \mathcal{S}^1$, wobei $\mathcal{S}^1 := \{x \in \mathbb{R}^2 : \|x\|_2 = 1\}$ die Einheitssphäre im Raum \mathbb{R}^2 bezeichnet. Besonders der Zusammenhang zu Orthogonalprojektoren auf die Unterräume, und damit die Aussage von Satz 3.3, ist zu erkennen:

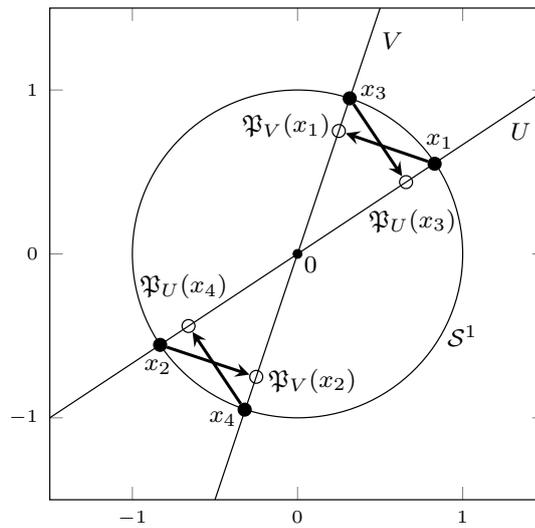


Abbildung 3.1: Veranschaulichung des Abstandes zweier Unterräume (Definition 3.1)

Der Abstand von U und V ergibt sich in der obigen Abbildung als

$$\text{dist}(U, V) = \|x_1 - \mathfrak{P}_V(x_1)\| = \|x_2 - \mathfrak{P}_V(x_2)\| = \|x_3 - \mathfrak{P}_U(x_3)\| = \|x_4 - \mathfrak{P}_U(x_4)\|.$$

◁

Wir wollen Satz 3.3 nun auf den Abstand der durch C und \hat{C} definierten aktiven Unterräume übertragen. Die betrachteten Unterräume sind $\text{im}(W_1)$ und $\text{im}(\hat{W}_1)$, die entsprechenden Orthogonalprojektoren sind $\mathfrak{P}_{\text{im}(W_1)} = W_1 W_1^T$ sowie $\mathfrak{P}_{\text{im}(\hat{W}_1)} = \hat{W}_1 \hat{W}_1^T$ (siehe Bemerkung 2.2). Mit Satz 3.3 gilt

$$\text{dist}\left(\text{im}(W_1), \text{im}(\hat{W}_1)\right) = \|W_1 W_1^T - \hat{W}_1 \hat{W}_1^T\|.$$

Um den Abstand der aktiven Unterräume von C und \hat{C} abzuschätzen, nutzen wir eine Aussage von Golub und Van Loan [GVL96] und kombinieren diese mit Korollar 3.8 von Constantine [Con15]. Zunächst sei allerdings Theorem 3.7 von Constantine wiedergegeben. Dieses gibt eine Abschätzung für den Abstand zwischen C und \hat{C} an und ist wesentlich für den Beweis von Korollar 3.5:

3 Umsetzung des Verfahrens

Satz 3.4 ([Con15], Theorem 3.7). *Es sei $\|\nabla_x f\| \leq L$ für alle $x \in \mathcal{X}$. Dann gilt für alle $\varepsilon \in (0, 1]$, dass*

$$\mathbb{P} \left[\|\hat{C} - C\| \geq \varepsilon \|C\| \right] \leq 2m \exp \left(\frac{-3M\lambda_1\varepsilon^2}{8L^2} \right).$$

Der Beweis nutzt eine ähnliche Bernstein-Ungleichung wie in Satz 3.1 und findet sich in [Con15]. Ähnlich zu Korollar 3.2 folgt auch aus diesem Satz eine Aussage zur Anzahl der Samples die notwendig sind, um eine relative Genauigkeit von ε zu erreichen.

Korollar 3.5 ([Con15], Korollar 3.8). *Für alle $\varepsilon \in (0, 1]$ folgt aus*

$$M = \Omega \left(\frac{L^2}{\lambda_1\varepsilon^2} \log(m) \right),$$

dass $\|\hat{C} - C\| \leq \varepsilon \|C\|$ mit hoher Wahrscheinlichkeit gilt.

Beweis. Wir nutzen Satz 3.4 und wählen M als

$$M \geq \frac{8L^2}{3\lambda_1\varepsilon^2} (\beta + 1) \log(m)$$

mit $\beta > 0$. Dann folgt

$$\begin{aligned} \mathbb{P} \left[\|\hat{C} - C\| \geq \varepsilon \|C\| \right] &\leq 2m \exp \left(\frac{-3M\lambda_1\varepsilon^2}{8L^2} \right) \\ &\leq 2m \exp \left(-(\beta + 1) \log(m) \right) \\ &= 2m^{-\beta} \end{aligned}$$

und damit die Behauptung, denn es gilt

$$\mathbb{P} \left[\|\hat{C} - C\| \leq \varepsilon \|C\| \right] \geq 1 - \mathbb{P} \left[\|\hat{C} - C\| \geq \varepsilon \|C\| \right] \geq 1 - 2m^{-\beta} \longrightarrow 1$$

für $\beta \longrightarrow \infty$. ■

Wir wollen dieses Korollar mit einem Lemma von Golub und Van Loan kombinieren. Um den Abstand der Unterräume $\text{im}(W_1)$ und $\text{im}(\hat{W}_1)$ abschätzen zu können, wird eine Verbindung zu der Norm von $E = \hat{C} - C$ hergestellt.

Lemma 3.6 ([GVL96], Korollar 8.1.11). *Es seien C und E symmetrische $m \times m$ -Matrizen sowie $\hat{C} = C + E$. Die Eigenwerte von C und \hat{C} seien $\lambda_1, \dots, \lambda_m$ und $\hat{\lambda}_1, \dots, \hat{\lambda}_m$ mit den Eigenvektormatrizen*

$$W = \begin{bmatrix} W_1 & W_2 \end{bmatrix} \quad \text{und} \quad \hat{W} = \begin{bmatrix} \hat{W}_1 & \hat{W}_2 \end{bmatrix},$$

wobei W_1 und \hat{W}_1 die ersten $n < m$ Spalten enthalten. Falls $\lambda_n > \lambda_{n+1}$ und $\|E\| \leq \frac{\lambda_n - \lambda_{n+1}}{5}$, dann gilt

$$\text{dist} \left(\text{im}(W_1), \text{im}(\hat{W}_1) \right) \leq \frac{4\|W_2^T E W_1\|}{\lambda_n - \lambda_{n+1}}.$$

Wir erhalten als zentrales Ergebnis dieses Abschnitts folgendes Korollar, das den Abstand der durch W_1 und \hat{W}_1 definierten Unterräume in Abhängigkeit von $\varepsilon > 0$, und damit der Wahl von M , beschreibt.

Korollar 3.7 ([Con15], Korollar 3.10). *Es sei $\varepsilon > 0$ sodass $\varepsilon \leq \frac{\lambda_n - \lambda_{n+1}}{5\lambda_1}$ und es sei M wie in Korollar 3.5 gewählt. Dann gilt*

$$\text{dist}\left(\text{im}(W_1), \text{im}(\hat{W}_1)\right) \leq \frac{4\lambda_1\varepsilon}{\lambda_n - \lambda_{n+1}} \quad (3.2)$$

mit hoher Wahrscheinlichkeit.

Beweis. Es sei $E = \hat{C} - C$, dann gilt mit den Voraussetzungen an ε und M sowie Korollar 3.5 mit hoher Wahrscheinlichkeit

$$\|E\| \leq \varepsilon\|C\| = \varepsilon\lambda_1 \leq \lambda_1 \frac{\lambda_n - \lambda_{n+1}}{5\lambda_1} = \frac{\lambda_n - \lambda_{n+1}}{5}.$$

Dies entspricht gerade der Voraussetzung an $\|E\|$ aus Lemma 3.6. Bei obiger Abschätzung wurde ausgenutzt, dass $\|C\| = \lambda_1$ gilt, denn Eigenwerte und Singulärwerte der symmetrischen und positiv-semidefiniten Matrix C stimmen überein (siehe auch Abschnitt 3.3). Die Spektralnorm ist submultiplikativ und selbstadjungiert, für eine orthogonale Matrix $A \in \mathbb{R}^{m \times n}$ gilt zudem $\|A\| = 1$. Somit ist

$$\|W_2^T E W_1\| \leq \|W_2^T\| \|E\| \|W_1\| = \underbrace{\|W_2\|}_{=1} \|E\| \underbrace{\|W_1\|}_{=1} = \|E\|,$$

denn W_1 und W_2 sind jeweils orthogonal. Aus Lemma 3.6 und der Abschätzung $\|E\| \leq \varepsilon\lambda_1$ folgt

$$\text{dist}\left(\text{im}(W_1), \text{im}(\hat{W}_1)\right) \leq \frac{4\|W_2^T E W_1\|}{\lambda_n - \lambda_{n+1}} \leq \frac{4\|E\|}{\lambda_n - \lambda_{n+1}} \leq \frac{4\lambda_1\varepsilon}{\lambda_n - \lambda_{n+1}}$$

mit hoher Wahrscheinlichkeit und damit die Behauptung. ■

Der Abstand zwischen dem mittels \hat{C} approximierten aktiven Unterraum und dem durch C definierten Unterraum hängt also insbesondere von dem Wert $\lambda_n - \lambda_{n+1}$ ab. Dieser Zusammenhang erklärt die Wahl der Dimension n des aktiven Unterraums in Algorithmus 1 beziehungsweise Gleichung (2.5). Je größer die Differenz von λ_n und λ_{n+1} , desto besser ist die Abschätzung (3.2). Zudem ist die Wahl von ε und damit die Größe von M für den Abstand von entscheidender Bedeutung. In der Praxis sind die Werte L und λ_1 meistens jedoch nicht bekannt. Daher kann M auch nicht ohne Weiteres entsprechend der Vorgabe aus Korollar 3.2 oder Korollar 3.5 gewählt werden. Constantine schlägt in [Con15] vor,

$$M \geq \alpha k \log(m)$$

Samples in Algorithmus 2 zu nutzen. Hierbei stellt α einen Oversampling-Faktor dar, der zwischen 2 und 10 gewählt wird, und k entspricht der größten handhabbaren Dimension, auf die reduziert werden kann.

3.2 Alternativen zur Verwendung von Gradienten

In vielen Anwendungen trifft die Annahme, dass wir Gradienten exakt berechnen können, nicht zu. Daher wollen wir in diesem Abschnitt das Verfahren aus Algorithmus 2 auf den Fall verallgemeinern, dass uns kein Black-Box-Algorithmus zur exakten Berechnung von

Gradienten zur Verfügung steht. Stattdessen gehen wir davon aus, dass sich ausschließlich einzelne Funktionswerte berechnen lassen.

3.2.1 Approximation von Gradienten

Ein naheliegender Ansatz wäre die Approximation von Gradienten durch die näherungsweise Bestimmung sämtlicher Richtungsableitungen entlang der m Koordinatenrichtungen. Dies führt zu mehrfachem Auswerten der Funktion f an bestimmten Punkten um die M zufällig gewählten Punkte herum. Beispielsweise mittels finiter Differenzen lässt sich so der Gradient approximativ angeben. Auch die allgemeine Annahme, dass der Gradient mit einem gewissen Fehler berechnet wird, führt zur Konvergenz des Verfahrens, wenn der Fehler mit wachsender Genauigkeit verschwindet (siehe Lemma 3.11, Theorem 3.12 und Theorem 3.13 in [Con15]). Wir wollen an dieser Stelle allerdings Ansätze betrachten, deren Rechenaufwand im Allgemeinen deutlich geringer ist, als die Approximation von Gradienten mittels finiter Differenzen oder ähnlicher Verfahren. Dies ermöglicht eine praktikablere Berechnung aktiver Unterräume.

3.2.2 Lokale lineare Modelle

Die Approximation von Gradienten mit Hilfe der hier betrachteten linearen Modelle beruht auf einer Linearisierung der Zielfunktion f in einem gewissen Bereich um den betrachteten Punkt. Ein wesentlicher Unterschied zu anderen Verfahren ist hierbei, dass die Punkte, an denen der Funktionswert bestimmt werden muss, nicht deterministisch gewählt werden. Zudem ist es bei diesem Ansatz vorgesehen, dass ein einzelner berechneter Funktionswert zur Bestimmung mehrerer Gradienten beiträgt. In diesem Abschnitt sollen zunächst lokale lineare Modelle diskutiert werden, der nächste Abschnitt behandelt den Spezialfall des globalen linearen Modells.

Der Algorithmus sieht wie folgt aus:

Algorithmus 3 Aktive Unterräume (praktisch, lokale lineare Modelle)

- 1: Ziehe N unabhängige Samples $\{x_j\}_{j=1}^N$ bezüglich der Dichte ρ
- 2: Berechne $q_j := f(x_j)$ für jedes $j = 1, \dots, N$
- 3: Ziehe M unabhängige Samples $\{x'_i\}_{i=1}^M$ bezüglich der Dichte ρ
- 4: **for** $i = 1, \dots, M$ **do**
- 5: Finde diejenigen p Punkte aus der Menge $\{x_j\}$, die am nächsten zu x'_i liegen
- 6: Bezeichne die Menge der zugehörigen Indizes als D_i , das heißt $D_i \subseteq \{1, \dots, N\}$
- 7: Nutze die Methode der kleinsten Quadrate, um mittels linearer Regression die Koeffizienten $c_i \in \mathbb{R}$ und $b_i \in \mathbb{R}^m$ zu berechnen mit

$$q_j \approx c_i + b_i^T x_j, \quad j \in D_i$$

8: **end for**

9: Approximiere C durch $\hat{C} := \frac{1}{M} \sum_{i=1}^M b_i b_i^T$

10: Verfahre wie in Algorithmus 2 in den Schritten 4-6

Das Verfahren approximiert an den Punkten $x'_i \in \mathcal{X}$, $i = 1, \dots, M$, den Gradienten von f mittels linearer Regression, der Funktionswert von f im Punkt x'_i muss hierzu nicht bestimmt werden. Zu diesem Zweck wird eine Menge von Wertepaaren (x_j, q_j) für $j = 1, \dots, N$ berechnet, aus der jeweils p Punkte zur Definition des linearen Ausgleichsproblems verwendet werden. Damit kann ein Wertepaar bei der Approximation mehrerer Gradienten eingesetzt werden. Die folgende Abbildung veranschaulicht die Wahl der Wertepaare für die Konstruktion der linearen Modelle. Die kleineren blauen Punkte stellen die Menge der N Samples $\{x_j\}$ dar, die größeren roten Punkte markieren die M Stellen $\{x'_j\}$, an denen der Gradient approximiert wird. In diesem Beispiel wurde $p = 3$ gewählt.

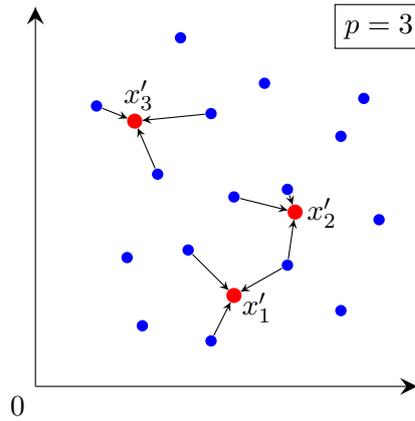


Abbildung 3.2: Grafische Veranschaulichung der Konstruktion lokaler linearer Modelle zur Approximation von Gradienten

Die Anzahl M der näherungsweise bestimmten Gradienten sollte wie in Algorithmus 2 als $M = \alpha k \log(m)$ gewählt werden, für die Anzahl N der Wertepaare, die berechnet werden, sollte $N \geq \alpha m$ gelten (siehe [Con15], Abschnitt 1.1). Der Wert p wird durch $m + 1 \leq p \leq N$ eingeschränkt, damit das lineare Ausgleichsproblem in Schritt 7 von Algorithmus 3 wohldefiniert und eindeutig lösbar ist. Das Ausgleichsproblem ergibt sich für ein $i \in \{1, \dots, M\}$ als

$$\min_{\substack{c_i \in \mathbb{R} \\ b_i \in \mathbb{R}^m}} \left\| \begin{bmatrix} \left[\begin{array}{c} 1 \\ \vdots \\ 1 \end{array} \right] \Psi \\ \left[\begin{array}{c} c_i \\ b_i \end{array} \right] \end{bmatrix} - \left[\begin{array}{c} q \\ \end{array} \right] \right\|_2,$$

wobei die Zeilen von $\Psi \in \mathbb{R}^{p \times m}$ den Vektoren x_j^T mit $j \in D_i$ entsprechen und $q \in \mathbb{R}^p$ die zugehörigen Funktionswerte q_j enthält. Zur Lösung dieses Minimierungsproblems kann beispielsweise die Moore-Penrose-Inverse eingesetzt werden. Die GNU Scientific Library [GTJ⁺21] für numerische Berechnungen in C++ stellt für die Lösung eines solchen Ausgleichsproblems die Funktion `gsl_linalg_SV_solve` zur Verfügung. Nachdem in den Schritten 4-8 einige Gradienten von f approximiert wurden, bestimmen wir die Matrix \hat{C} wie in Algorithmus 2, allerdings unter Verwendung der durch lineare Modelle approximierten Gradienten. Die Berechnung der aktiven Unterräume aus der Eigenwertzerlegung von \hat{C} folgt anschließend dem aus Algorithmus 2 bekannten Schema.

3.2.3 Globale lineare Modelle

Wird Algorithmus 3 für den Fall $p = N$ betrachtet, so lässt sich das Vorgehen weiter vereinfachen. Dann gilt nämlich $D_i = \{x_j\}_{j=1}^N$ für alle $i = 1, \dots, M$ und damit $b_i = b_{i'}$ für alle $i, i' = 1, \dots, M$. Die Matrix \hat{C} vereinfacht sich zu

$$\hat{C} = b \cdot b^T,$$

wobei $b = b_i$ für $i \in \{1, \dots, M\}$ beliebig sei. Wir gehen davon aus, dass $b \neq 0$ gilt, ansonsten gibt es keinen aktiven Unterraum denn f ist auf \mathcal{X} konstant oder aber der aktive Unterraum konnte aus den gegebenen Samples nicht berechnet werden. Damit ist \hat{C} eine Matrix mit Rang 1, das heißt insbesondere, dass $m - 1$ Eigenwerte von \hat{C} bereits null sind. Der einzige positive Eigenwert ist $b^T b$, der normierte Eigenvektor zu $\lambda_1 = b^T b$ ist $\hat{w} = b / \|b\|$.

Das soeben beschriebene Verfahren nutzt ein globales lineares Modell für den Gradienten, der Gradient wird also in jedem Punkt x'_i durch denselben Vektor b approximiert. Der Rechenaufwand reduziert sich auf die Berechnung der Funktionswerte sowie eine einzige lineare Regression, um den Vektor b zu bestimmen. Allerdings liefert der Ansatz in jedem Fall einen eindimensionalen aktiven Unterraum, auch wenn dieser nicht vorhanden ist. Trotzdem ist es sinnvoll, das Verfahren als ersten Ansatz zur Dimensionsreduktion vor jeglicher Parameterstudie in Betracht zu ziehen. Beschreibt der berechnete eindimensionale aktive Unterraum das Verhalten der Funktion f bereits hinreichend gut, so kann der Rechenaufwand in den Parameterstudien häufig erheblich verringert werden. Ist der eindimensionale Unterraum hingegen nicht ausreichend, so können anschließend lokale lineare Modelle für den Gradienten eingesetzt werden. In Kapitel 4 werden wir Beispiele betrachten, in denen ein globales lineares Modell zum Einsatz kommt. Zudem werden wir eine grafische Möglichkeit kennenlernen, mit deren Hilfe sich einschätzen lässt, ob tatsächlich ein eindimensionaler aktiver Unterraum vorliegt (siehe insbesondere Abschnitt 4.1.3).

Zum Abschluss des Abschnitts sei im Folgenden der Algorithmus zur Berechnung eines eindimensionalen aktiven Unterrums \hat{w} unter Verwendung eines globalen linearen Modells für den Gradienten zusammengefasst:

Algorithmus 4 Aktive Unterräume (praktisch, globales lineares Modell)

- 1: Ziehe N unabhängige Samples $\{x_j\}_{j=1}^N$ bezüglich der Dichte ρ
- 2: Berechne $q_j := f(x_j)$ für jedes $j = 1, \dots, N$
- 3: Nutze die Methode der kleinsten Quadrate, um mittels linearer Regression die Koeffizienten $c \in \mathbb{R}$ und $b \in \mathbb{R}^m$ zu berechnen mit

$$q_j \approx c + b^T x_j, \quad j = 1, \dots, N$$

- 4: Normiere den Vektor b und setze

$$\hat{w} = b / \|b\|$$

3.3 Eigenwertzerlegung der Matrix C

Wir sind in Abschnitt 2.1 und auch in Algorithmus 2 davon ausgegangen, dass sich die Eigenwertzerlegung der Matrix C beziehungsweise der Matrix \hat{C} exakt berechnen lässt. In diesem Abschnitt wollen wir kurz beschreiben, wie dies praktisch umgesetzt werden kann. Die Eigenschaften aus Lemma 2.1, dass die Matrix C symmetrisch und positiv-semidefinit ist, übertragen sich auch auf die Matrix \hat{C} wie sie in Algorithmus 2 und Algorithmus 3 berechnet wurde. Es lässt sich insbesondere folgende allgemeine Aussage über Matrizen dieser Form zeigen:

Lemma 3.8. *Es seien $M \in \mathbb{N}$ und $a_i \in \mathbb{R}^m$ für $i = 1, \dots, M$. Dann ist die Matrix*

$$\hat{C} := \frac{1}{M} \sum_{i=1}^M a_i a_i^T$$

symmetrisch und positiv-semidefinit.

Beweis. Für eine Matrix \hat{C} der obigen Form gilt

$$\hat{C}^T = \left(\frac{1}{M} \sum_{i=1}^M a_i a_i^T \right)^T = \frac{1}{M} \sum_{i=1}^M (a_i a_i^T)^T = \frac{1}{M} \sum_{i=1}^M a_i a_i^T = \hat{C},$$

das heißt \hat{C} ist symmetrisch. Zudem gilt für alle $w \in \mathbb{R}^m$, dass

$$w^T \hat{C} w = w^T \left(\frac{1}{M} \sum_{i=1}^M a_i a_i^T \right) w = \frac{1}{M} \sum_{i=1}^M w^T a_i a_i^T w = \frac{1}{M} \sum_{i=1}^M (w^T a_i)^2 \geq 0,$$

also ist \hat{C} positiv-semidefinit. ■

Im Folgenden sei A eine beliebige symmetrische und positiv-semidefinite $m \times m$ -Matrix. Nach Lemma 2.1 und Lemma 3.8 lassen sich die folgenden Ergebnisse auch auf die Matrix C aus der theoretischen Berechnung aktiver Unterräume sowie die Matrizen \hat{C} aus den vorangehenden beiden Abschnitten übertragen.

Das Ziel ist es nun, eine Eigenwertzerlegung der Matrix A zu berechnen. Dabei wollen wir auf eine Singulärwertzerlegung von A zurückgreifen. Diese lässt sich beispielsweise mit Householder- oder Givens-Transformationen numerisch stabil berechnen. Wir betrachten folgenden Satz von Golub und Van Loan, der die Schur-Zerlegung im Falle einer symmetrischen Matrix beschreibt:

Satz 3.9 ([GVL96], Theorem 8.1.1 (Symmetric Schur Decomposition)). *Es sei $A \in \mathbb{R}^{m \times m}$ symmetrisch, dann existiert eine reelle orthogonale Matrix $W \in \mathbb{R}^{m \times m}$, sodass*

$$W^T A W = \Lambda = \text{diag}(\lambda_1, \dots, \lambda_m).$$

Des Weiteren gilt $Aw_k = \lambda_k w_k$ für alle $k = 1, \dots, m$, wobei $w_k \in \mathbb{R}^m$ die k -te Spalte der Matrix W bezeichnet.

Wir nehmen zusätzlich zur Symmetrie an, dass A positiv-semidefinit ist, es gilt daher $\lambda_k \geq 0$ für alle $k = 1, \dots, m$. Somit liefert die Schur-Zerlegung $W^T A W = \Lambda$ durch Um-

stellen insbesondere eine Singulärwertzerlegung $A = W\Lambda W^T$, denn es gilt $WW^T = I_m$. Hierbei haben wir angenommen, dass die Spalten der Matrix W so sortiert sind, dass $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m \geq 0$ gilt. Die Eigenwerte und Singulärwerte von A stimmen also überein. Die Schur-Zerlegung entspricht nun der Eigenwertzerlegung von A , denn W ist eine orthogonale Matrix, deren Spalten die Eigenvektoren von A bilden, und Λ ist die Diagonalmatrix der absteigend sortierten Eigenwerte. Es genügt also, eine Singulärwertzerlegung von A zu berechnen, um gleichzeitig eine Eigenwertzerlegung zu erhalten, denn die Singulärwertzerlegung ist bis auf das Vorzeichen der Singulärvektoren (in diesem Fall der Eigenvektoren) eindeutig festgelegt.

Die Bestimmung der Eigenwertzerlegung in Algorithmus 2 und Algorithmus 3 reduziert sich somit auf die Berechnung der Singulärwertzerlegung der Matrix \hat{C} . In C++ lässt sich dies beispielsweise mit der Funktion `gsl_linalg_SV_decomp` aus der GNU Scientific Library [GTJ⁺21] bewerkstelligen.

3.4 Kombiniertes aktiver Unterraum zu mehreren Funktionen

Wir haben uns bis dato ausschließlich darauf fokussiert, aktive Unterräume einer einzelnen Funktion $f: \mathcal{X} \rightarrow \mathbb{R}$ zu berechnen. In diesem Abschnitt soll ein Ansatz erläutert werden, wie sich mehrere Funktionen $f_i: \mathcal{X} \rightarrow \mathbb{R}$ für $i = 1, \dots, p$ mit $p \in \mathbb{N}$ behandeln lassen. Dabei soll nur ein aktiver Unterraum bestimmt werden, der für alle Funktionen in gewissem Sinne geeignet ist. Ein Beispiel für die Anwendung des hier vorgestellten Verfahrens findet sich in Abschnitt 4.1.5.

Wir beschränken uns zunächst auf den Fall, dass sämtliche Funktionen f_1, \dots, f_p einen eindimensionalen aktiven Unterraum besitzen. Dieser kann beispielsweise mit Hilfe eines globalen linearen Modells berechnet werden, siehe Abschnitt 3.2.3. Ji et al. präsentieren in [JWZ⁺18] ein Verfahren, um sogenannte *shared subspaces* zu berechnen. Diese Unterräume sind so gewählt, dass sich die Dichten der Funktionswerte aller Funktionen f_1, \dots, f_p gleichzeitig approximieren lassen. Shared subspaces stellen jedoch im Allgemeinen keine Richtungen dar, sie müssen nicht normiert sein. Wir werden daher an dieser Stelle einen etwas anderen Ansatz verfolgen mit dem Ziel, die aktiven Unterräume von f_1, \dots, f_p durch eine einzige Richtung anzunähern.

Es seien w_1, \dots, w_p die eindimensionalen aktiven Unterräume der Funktionen f_1, \dots, f_p . Die einzelnen Funktionen werden durch $\alpha_i \in \mathbb{R}_+$ gewichtet. Wird $\alpha_{i'}$ für $i' \in \{1, \dots, p\}$ im Vergleich zu den Gewichten der anderen Funktionen groß gewählt, so liegt ein besonderer Fokus auf einer möglichst guten Approximation des aktiven Unterraums der Funktion $f_{i'}$. Der eindimensionale aktive Unterraum der nun konstruiert werden soll, kombiniert die aktiven Unterräume w_1, \dots, w_p . Dazu betrachten wir diejenige Richtung, die die Summe der gewichteten Abstandsquadrate zu den Richtungen w_1, \dots, w_p minimiert. Wir suchen also $v^* \in \mathbb{R}^m$ mit $\|v^*\|_2 = 1$ und

$$v^* = \operatorname{argmin}_{\substack{v \in \mathbb{R}^m \\ \|v\|_2 = 1}} \sum_{i=1}^p \alpha_i \|v - w_i\|_2^2. \quad (3.3)$$

Wir nutzen nun folgenden Satz, der v^* in einem allgemeineren Kontext beschreibt:

Satz 3.10. *Es sei V ein endlich-dimensionaler Vektorraum über den reellen Zahlen und $\|\cdot\|_V$ eine durch ein Skalarprodukt induzierte Norm auf V . Zudem seien $r \in \mathbb{R}_+$ sowie $w_1, \dots, w_p \in V$ und $\alpha_1, \dots, \alpha_p \in \mathbb{R}$ mit $p \in \mathbb{N}$ und $\sum_{i=1}^p \alpha_i w_i \neq 0$. Des Weiteren sei*

$$v^* := \operatorname{argmin}_{\substack{v \in V \\ \|v\|_V = r}} \sum_{i=1}^p \alpha_i \|v - w_i\|_V^2.$$

Dann gilt

$$v^* = \frac{r \cdot \sum_{i=1}^p \alpha_i w_i}{\left\| \sum_{i=1}^p \alpha_i w_i \right\|_V}, \quad (3.4)$$

insbesondere ist v^* eindeutig.

Beweis. Die Norm $\|\cdot\|_V$ werde durch das Skalarprodukt $\langle \cdot, \cdot \rangle_V$ induziert, das heißt $\|v\|_V = \sqrt{\langle v, v \rangle_V}$ für alle $v \in V$. Dann gilt für alle $i = 1, \dots, p$ und für alle $v \in V$ mit $\|v\|_V = r$, dass

$$\begin{aligned} \|v - w_i\|_V^2 &= \langle v - w_i, v - w_i \rangle_V \\ &= \underbrace{\langle v, v \rangle_V}_{=r} - 2\langle v, w_i \rangle_V + \langle w_i, w_i \rangle_V \\ &= r - 2\langle v, w_i \rangle_V + \langle w_i, w_i \rangle_V. \end{aligned}$$

Damit folgt

$$\begin{aligned} v^* &= \operatorname{argmin}_{\substack{v \in V \\ \|v\|_V = r}} \sum_{i=1}^p \alpha_i \|v - w_i\|_V^2 \\ &= \operatorname{argmin}_{\substack{v \in V \\ \|v\|_V = r}} \sum_{i=1}^p \alpha_i \left(r - 2\langle v, w_i \rangle_V + \langle w_i, w_i \rangle_V \right) \\ &= \operatorname{argmin}_{\substack{v \in V \\ \|v\|_V = r}} \left(\sum_{i=1}^p \alpha_i r - 2 \sum_{i=1}^p \alpha_i \langle v, w_i \rangle_V + \sum_{i=1}^p \alpha_i \langle w_i, w_i \rangle_V \right). \end{aligned}$$

Die erste und die letzte Summe sind unabhängig von v und damit unerheblich für das sich ergebende Minimum. Somit vereinfacht sich die obige Formel für v^* zu

$$\begin{aligned} v^* &= \operatorname{argmin}_{\substack{v \in V \\ \|v\|_V = r}} \left(-2 \sum_{i=1}^p \alpha_i \langle v, w_i \rangle_V \right) \\ &= \operatorname{argmax}_{\substack{v \in V \\ \|v\|_V = r}} \sum_{i=1}^p \alpha_i \langle v, w_i \rangle_V \\ &= \operatorname{argmax}_{\substack{v \in V \\ \|v\|_V = r}} \left\langle v, \sum_{i=1}^p \alpha_i w_i \right\rangle_V. \end{aligned}$$

3 Umsetzung des Verfahrens

Es sei

$$w := \sum_{i=1}^p \alpha_i w_i,$$

nach Voraussetzung gilt $w \neq 0$ also auch $\|w\|_V \neq 0$. Mit der Cauchy-Schwarz-Ungleichung gilt für alle $v \in V$ mit $\|v\|_V = r$ die Ungleichung

$$\langle v, w \rangle_V \leq \underbrace{\|v\|_V}_{=r} \cdot \|w\|_V = r \cdot \|w\|_V,$$

wobei Gleichheit genau für $v = \lambda w$ mit $\lambda \in \mathbb{R}$ eintritt. Damit gilt insbesondere

$$\max_{\substack{v \in V \\ \|v\|_V = r}} \langle v, w \rangle_V \leq r \cdot \|w\|_V.$$

Um die Bedingung $\|v^*\|_V = r$ zu sichern, muss $\lambda = r/\|w\|_V$ oder $\lambda = -r/\|w\|_V$ gewählt werden und $v^* = \lambda w$.

Für den Fall $v^* = -r \cdot w / \|w\|_V$ gilt allerdings

$$\langle v^*, w \rangle_V = -\frac{r}{\|w\|_V} \langle w, w \rangle_V < 0 < r \cdot \|w\|_V.$$

Im Fall $v^* = r \cdot w / \|w\|_V$ folgt hingegen

$$\max_{\substack{v \in V \\ \|v\|_V = r}} \langle v, w \rangle_V \leq r \cdot \|w\|_V = \frac{r \cdot \|w\|_V^2}{\|w\|_V} = \frac{r}{\|w\|_V} \langle w, w \rangle_V = \left\langle \frac{r \cdot w}{\|w\|_V}, w \right\rangle_V = \langle v^*, w \rangle_V.$$

Damit gilt

$$\operatorname{argmax}_{\substack{v \in V \\ \|v\|_V = r}} \langle v, w \rangle_V = \frac{r \cdot w}{\|w\|_V}.$$

und es folgt

$$v^* = \operatorname{argmin}_{\substack{v \in V \\ \|v\|_V = r}} \sum_{i=1}^p \alpha_i \|v - w_i\|_V^2 = \frac{r \cdot w}{\|w\|_V} = \frac{r \cdot \sum_{i=1}^p \alpha_i w_i}{\left\| \sum_{i=1}^p \alpha_i w_i \right\|_V}.$$

Der Vektorraum V ist nach Voraussetzung endlich-dimensional, das heißt die Sphäre $\mathcal{S}^r := \{v \in V : \|v\|_V = r\}$ ist kompakt. Die stetige Funktion $\sum_{i=1}^p \alpha_i \|v - w_i\|_V^2$ nimmt auf \mathcal{S}^r somit ein Minimum sowie ein Maximum an. Mit obiger Argumentation ist das Minimum v^* eindeutig. ■

Bemerkung 3.6. Der Schwerpunkt eines Dreiecks, beziehungsweise eines Simplex, ist analog zu der Wahl von v^* in (3.4) definiert, allerdings ohne den Normierungsterm und damit die Einschränkung auf die Sphäre \mathcal{S}^r . Zudem besitzt der Schwerpunkt eine ganz ähnliche Eigenschaft wie v^* in Satz 3.10. Auch er minimiert die Summe der Abstandsquadrate zu den Eckpunkten des betrachteten Dreiecks oder Simplex.

In Satz 3.10 werden jeweils die quadrierten Abstände betrachtet. Würde stattdessen die Summe der Abstände minimiert, so wäre beispielsweise für $V = \mathbb{R}^2$ mit der Euklidischen Norm, $r = 1$, $p = 2$ und $\alpha_1 = \alpha_2 = 1$ jeder der beiden Vektoren w_1 und w_2 optimal. Die beiden Vektoren würden also nicht kombiniert, um v^* zu berechnen. ◁

In unserer Anwendung bei aktiven Unterräumen ist $V = \mathbb{R}^m$ und der Vektor v^* soll eine Richtung definieren, das heißt, wir setzen $r = 1$. Die genutzte Norm ist die Euklidische Norm auf dem Vektorraum \mathbb{R}^m , induziert durch das Standardskalarprodukt. Zudem sind die Vektoren w_1, \dots, w_p normiert. Nach Satz 3.10 ergibt sich der Vektor v^* als normierte Summe der Vektoren w_1, \dots, w_p , vergleiche Abbildung 3.3. Insbesondere lässt sich v^* ohne großen Aufwand berechnen.

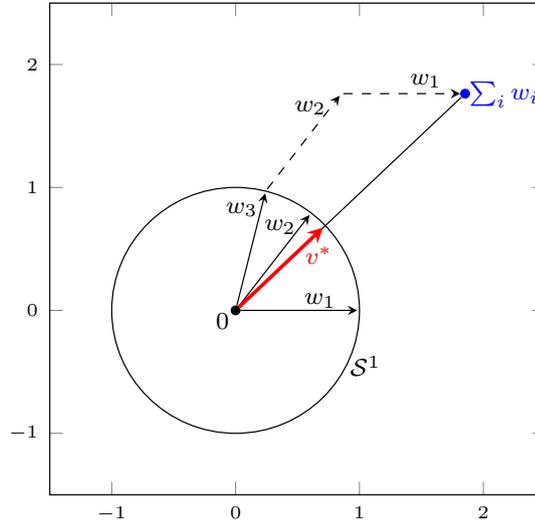


Abbildung 3.3: Beispielhafte Berechnung von v^* mit Satz 3.10

Das in diesem Abschnitt beschriebene Vorgehen kann allgemeiner aufgefasst werden. Statt dem eindimensionalen Wertebereich der reellen Zahlen, betrachten wir eine Funktion \tilde{f} mit mehrdimensionalem Wertebereich, das heißt

$$\tilde{f}: \mathcal{X} \rightarrow \mathbb{R}^p.$$

Die aktiven Unterräume könnten nun komponentenweise bestimmt und einzeln weiter verwendet werden. Mit v^* aus Satz 3.10 lässt sich allerdings auch ein eindimensionaler aktiver Unterraum berechnen, der in gewissem Maße zu allen Komponenten von \tilde{f} passt, sofern in den einzelnen Komponenten jeweils ein eindimensionaler aktiver Unterraum vorliegt.

An dieser Stelle bleibt die Frage offen, wie sich das Konzept der Berechnung von v^* auf den Fall übertragen lässt, dass kein eindimensionaler sondern ein mehrdimensionaler aktiver Unterraum in mindestens einer der Komponenten vorliegt. Dieses Problem ist ungleich komplizierter zu behandeln. Wir bezeichnen mit n_1, \dots, n_p die Dimensionen der aktiven Unterräume der einzelnen Komponenten von \tilde{f} . Zudem seien $W_1^1 = \begin{bmatrix} w_1^1 & \cdots & w_{n_1}^1 \end{bmatrix}, \dots, W_1^p = \begin{bmatrix} w_1^p & \cdots & w_{n_p}^p \end{bmatrix}$ die zugehörigen orthogonalen Matrizen, die die aktiven Unterräume bestimmen. Das Minimierungsproblem aus (3.3), das von v^* gelöst wird, kann in die Optimierungsaufgabe

$$V^* = \underset{V = [v_1, \dots, v_{n_{\max}}] \in \mathbb{V}_{n_{\max}}^m}{\operatorname{argmin}} \sum_{i=1}^p \sum_{j=1}^{n_i} \min_{k=1, \dots, n_{\max}} \|v_k - w_j^i\|_2^2$$

übertragen werden. Hierbei bezeichnet $\mathbb{V}_n^m = \{V \in \mathbb{R}^{m \times n} : V^T V = I_n\}$ die reelle Man-

3 Umsetzung des Verfahrens

nigfaltigkeit der spalten-orthogonalen Matrizen (die reelle Stiefel-Mannigfaltigkeit) und $n_{\max} = \max_{i=1, \dots, p} n_i$ sei die maximale Dimension eines aktiven Unterraums in einer der p Komponenten von \tilde{f} . Die Matrix V^* ist also eine orthogonale Matrix und definiert einen n_{\max} -dimensionalen aktiven Unterraum. Um V^* berechnen zu können, muss obiges Optimierungsproblem gelöst werden. Das größte Problem stellt hierbei die Bedingung dar, dass V^* orthogonal sein soll. Ansätze zur Lösung von Optimierungsaufgaben über der orthogonalen Gruppe finden sich in [SC13] sowie [YE03], diese müssten allerdings auf spaltenorthogonale, nicht-quadratische Matrizen übertragen werden, um hier Verwendung zu finden.

Kapitel 4

Berechnung aktiver Unterräume in praktischen Anwendungen

Das letzte Kapitel hat gezeigt, wie sich aktive Unterräume konkret berechnen lassen, ohne den Wert von Integralen exakt zu ermitteln. Insbesondere haben wir gesehen, wie sich das Vorgehen für den Fall anpassen lässt, dass der Gradient der untersuchten Funktion nicht zur Verfügung steht. In der Praxis wurden aktive Unterräume unter anderem beim Design von Strömungsmaschinen zur Umwandlung von Energie aus einem Fluid in kinetische Energie [BPDSC17] sowie zur Sensitivitätsanalyse bei wiederaufladbaren Lithium-Ionen-Batterien [CD17] und hydrologischen Modellen [JGCM16] eingesetzt. In den folgenden Abschnitten betrachten wir zwei weitere Beispiele, in denen sich aktive Unterräume berechnen lassen. Zunächst wird jeweils das zugrundeliegende Modell vorgestellt, anschließend werden die numerischen Ergebnisse näher beleuchtet. Der folgende Abschnitt beschäftigt sich mit der Simulation des Krankheitsverlaufs einer HIV-Erkrankung. Daraufhin werden wir eine Strömungssimulation aus der Aerodynamik studieren.

4.1 Simulation des Krankheitsverlaufs nach einer Infektion mit dem HI-Virus

Eine Infektion mit dem Humane Immundefizienz-Virus (HI-Virus) führt ohne entsprechende Behandlung nach mehreren Jahren zu einem Immunschwächesyndrom, dem sogenannten AIDS (Acquired Immune Deficiency Syndrome). Die Folgen der Krankheit sind in der Regel lebensbedrohlich [Mur12]. Um die Auswirkungen einer HIV-Infektion auf den menschlichen Körper quantifizieren zu können, wird die Zahl der sogenannten T-Helferzellen (auch $CD4^+$ -Lymphozyten genannt) betrachtet. Diese dienen dem HI-Virus als Wirtszelle. Das Virus vermehrt sich also mit Hilfe der T-Helferzellen und zerstört diese anschließend. Je geringer die Zahl der T-Helferzellen im Körper, desto stärker ist die Immunabwehr des Körpers geschwächt, denn die T-Helferzellen übernehmen wichtige Aufgaben bei der Steuerung des Immunsystems (siehe beispielsweise [Mur12] für weitere Informationen). Im folgenden Abschnitt wird ein Modell vorgestellt, mit dessen Hilfe sich der Verlauf einer HIV-Erkrankung simulieren lässt. Anschließend berechnen wir aktive Unterräume zu ei-

nem festen Zeitpunkt nach der Infektion. In Abschnitt 4.1.5 betrachten wir abschließend eine Verallgemeinerung des Vorgehens, indem wir mehrere Zeitpunkte simultan behandeln.

4.1.1 Berechnung der Zahl der T-Helferzellen

Die Zahl der T-Helferzellen im Körper lässt sich näherungsweise durch die Lösung eines Systems gewöhnlicher Differentialgleichungen erster Ordnung beschreiben. Wir betrachten hier ein Modell von Hadjiandreou et al. [HCV07], welches den Krankheitsverlauf mittels sieben Differentialgleichungen widerspiegelt. Neben der Entwicklung der T-Helferzellen wird unter anderem die Zahl der Viren und die Population der Makrophagen durch das Differentialgleichungssystem beschrieben. Des Weiteren enthält das Modell siebenundzwanzig Parameter, die beispielsweise die natürliche Sterberate oder die Infektionsrate der T-Helferzellen angeben. Für uns ist im Folgenden die Zahl der T-Helferzellen von Interesse. Diese sei mit $T(t, x)$ bezeichnet, wobei $t \in \mathbb{R}_+$ die Zeit in Tagen seit der Infektion und $x \in \mathbb{R}^{27}$ die zusätzlichen Parameter darstellen. Zunächst werden wir $T(t, x)$ für ein festes $t \in \mathbb{R}_+$ betrachten, das heißt wir setzen als Zielfunktion

$$f_t: \mathcal{X} \rightarrow \mathbb{R}, \quad f_t(x) := T(t, x),$$

für $x \in \mathcal{X} \subseteq \mathbb{R}^{27}$. Der Parameterraum ist ein Quader, die genauen Werte für die obere sowie die untere Grenze finden sich in [LP17]. Durch Skalierung und Verschiebung können wir annehmen, dass $\mathcal{X} = [-1, 1]^{27}$ gilt, der Einfluss einzelner Parameter auf das Verhalten der Funktion f lässt sich dadurch besser vergleichen. Ein zufällig gewähltes $x \in [-1, 1]^{27}$ kann in eine zulässige Wahl \tilde{x} der Parameter in dem Differentialgleichungssystem transformiert werden, indem

$$\tilde{x} = \frac{1}{2} (\text{diag}(u - l) \cdot x + (u + l))$$

gesetzt wird, wobei $l \in \mathbb{R}^{27}$ und $u \in \mathbb{R}^{27}$ die obere sowie untere Grenze aus [LP17] seien. Zudem bezeichnet $\text{diag}(r)$ für $r \in \mathbb{R}^m$ die Diagonalmatrix, deren Diagonalelemente den Elementen aus r entsprechen, das heißt $\text{diag}(r) := [d_{ij}]_{i,j=1}^m$ mit $d_{ii} = r_i$ für $i = 1, \dots, m$ und $d_{ij} = 0$ für $i, j = 1, \dots, m$ mit $i \neq j$. Als Wahrscheinlichkeitsdichte nutzen wir eine Gleichverteilung. Um ein Sample $x = [x_i]_{i=1}^{27}$ aus \mathcal{X} zu berechnen, wird x_i für $i = 1, \dots, 27$ jeweils uniform zufällig aus $[-1, 1]$ gewählt.

Zur numerischen Lösung des Differentialgleichungssystems wurde die Funktion `ode45` in OCTAVE [REw22] verwendet. Diese baut auf der Dormand–Prince Methode auf, welche zu den Runge-Kutta-Verfahren zählt. Die ursprüngliche Implementierung von Loudon und Pankavich (siehe [LP17]) wurde hier ein wenig abgewandelt, insbesondere ist der eingesetzte Löser für das Differentialgleichungssystem verschieden. In der folgenden Abbildung sind zehn Simulationen der Population von T-Helferzellen in Abhängigkeit von der Zeit seit der Infektion mit dem HI-Virus dargestellt. Jede Simulation wurde mit einer unterschiedlichen zufälligen Wahl der Parameter $x \in \mathcal{X}$ durchgeführt und die Lösung des Differentialgleichungssystems jeweils bis zum Zeitpunkt $t = 3500$ berechnet. Bei sämtlichen Simulationen wurden die Anfangswerte für die einzelnen Funktionen im Differentialgleichungssystem identisch gewählt, diese gehen also nicht in den Parameterraum ein.

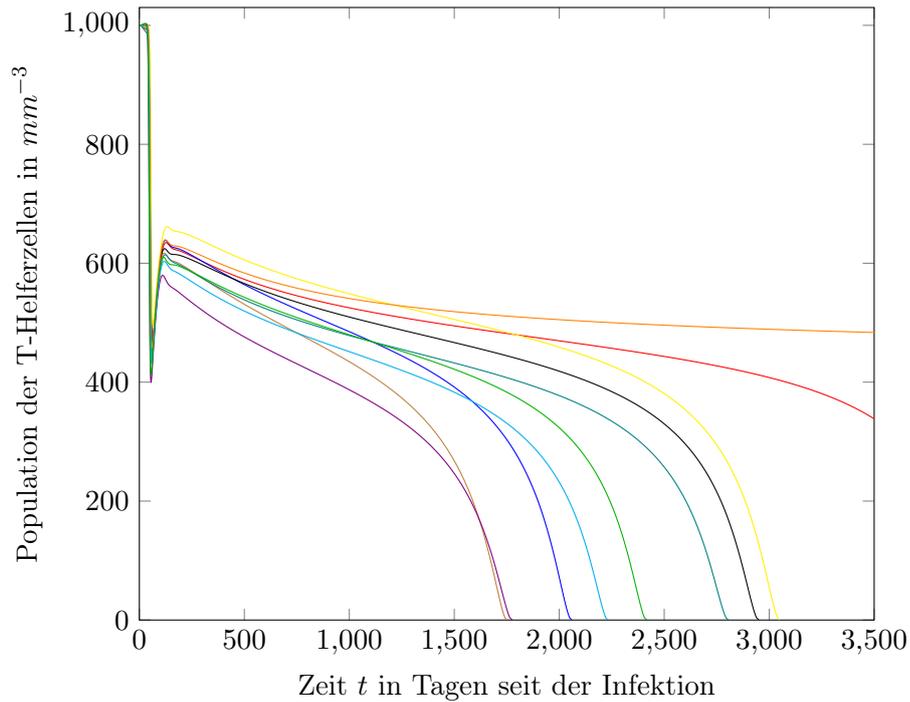


Abbildung 4.1: Simulationen der T-Helferzellen-Population mit unterschiedlichen Wahlen der 27 Parameter in Abhängigkeit von der Zeit t seit der Infektion

Die oben erkennbaren Krankheitsverläufe sind recht typisch für eine unbehandelte HIV-Erkrankung (vergleiche [Mur12]). Wenige Wochen nach der Infektion mit dem Virus verringert sich die Population der T-Helferzellen um knapp 60%, grippeähnliche Symptome treten ein. Diese halten für gewöhnlich einige Tage bis Wochen an, dann stabilisiert sich die Zahl der T-Helferzellen wieder. In der folgenden Latenzphase, die meist mehrere Jahre andauert, verringert sich die Zahl der T-Helferzellen stetig, allerdings leidet der Träger des Virus in dieser Phase nur unter geringen bis gar keinen körperlichen Beeinträchtigungen. Ungefähr zehn Jahre nach der Infektion sinkt die Zahl der T-Helferzellen in kurzer Zeit so rapide ab, dass es zu AIDS-definierenden Krankheitsbildern kommt.

Wir werden im folgenden Abschnitt die Population der T-Helferzellen zu einem festen Zeitpunkt in Abhängigkeit der Parameter $x \in \mathcal{X}$ betrachten und aktive Unterräume zu dieser Funktion berechnen.

4.1.2 Aktive Unterräume zu einem festen Zeitpunkt

Eine einzige Simulation, das heißt die Lösung des Differentialgleichungssystems für ein festes $x \in \mathcal{X}$ bis zum Zeitpunkt $t = 3500$, dauert knapp zehn Sekunden. Hierbei wurde OCTAVE auf einem Kern eines Intel[®] Core i5-2430M Prozessors mit 2,4 GHz Grundtakt verwendet. Dieses Beispiel passt somit zu der Annahme, dass die Berechnung eines einzelnen Funktionswertes bereits aufwändig ist.

Wir betrachten nun die Funktion $f_{1500}(x) = T(1500, x)$ mit $x \in \mathcal{X} = [-1, 1]^{27}$, das heißt die Anzahl der T-Helferzellen zum Zeitpunkt $t = 1500$ Tage nach der Infektion in Abhängigkeit von $x \in \mathcal{X}$. Loudon und Pankavich [LP17] approximieren den Gradienten von f_t mittels finiter Differenzen an 1000 zufällig gewählten Punkten (bezüglich der Gleich-

verteilung auf \mathcal{X}). Für die Approximation eines einzigen Gradienten, entsprechend dem in Abschnitt 3.1 in [LP17] beschriebenen Algorithmus, werden 28 Funktionsauswertungen benötigt, insgesamt müssen also 28000 Funktionswerte bestimmt werden. Statt finiter Differenzen verwenden wir hier lokale lineare Modelle für die Gradienten von f_{1500} wie in Abschnitt 3.2.2 beschrieben. Dabei nutzen wir insgesamt $N = 1000$ Funktionswerte, um an $M = 100$ Stellen den Gradienten mittels linearer Modelle basierend auf $p = 50$ Wertepaaren zu approximieren (vergleiche Algorithmus 3). Der Rechenaufwand beschränkt sich damit auf die Lösung von 1000 Differentialgleichungssystemen bis zum Zeitpunkt $t = 1500$. Wir werden sehen, dass die Ergebnisse, trotz des deutlich geringeren Rechenaufwands, denen von Loudon und Pankavich sehr ähneln und ebenfalls eine Struktur der Funktion erkennen lassen, die eine Dimensionsreduktion des Parameterraums ermöglicht.

Mit Hilfe der linearen Modelle konnte die Matrix C durch eine Approximation \hat{C} angenähert werden. Anschließend wurde die Eigenwertzerlegung $\hat{C} = \hat{W}\hat{\Lambda}\hat{W}^T$ berechnet. Die Eigenwerte der Matrix \hat{C} sind in der folgenden Grafik dargestellt.

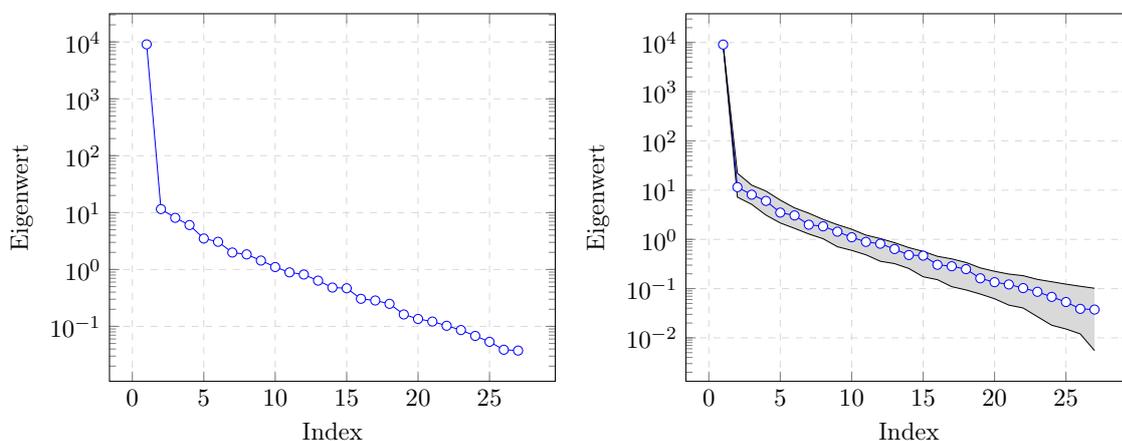


Abbildung 4.2: Eigenwerte der Matrix \hat{C} für den Zeitpunkt $t = 1500$ in absteigender Sortierung (links) und zusätzlich mit Bootstrap Intervallen (rechts)

In Abschnitt 2.2.2 haben wir gesehen, dass sich die Dimension des aktiven Unterraums durch die größte Lücke in den Eigenwerten bestimmen lässt. Die obige Abbildung 4.2 legt daher nahe, einen eindimensionalen aktiven Unterraum zu wählen, denn die Differenz zwischen dem ersten und zweiten Eigenwert ist größer als die paarweisen Abstände zwischen den restlichen Eigenwerten. Bezeichnen $\hat{w}_1, \dots, \hat{w}_{27}$ die orthonormierten Eigenvektoren von \hat{C} , also die Spalten von \hat{W} , so definiert \hat{w}_1 den aktiven Unterraum und $\hat{w}_2, \dots, \hat{w}_{27}$ erzeugen den inaktiven Unterraum.

Bemerkung 4.1. Die berechneten Eigenwerte hängen in lokalen linearen Modellen von der Wahl der Punkte ab, an denen der Gradient approximiert wird. Um die Variabilität der Eigenwerte, also den Einfluss der zufällig gewählten Punkte, abschätzen zu können, werden häufig sogenannte Bootstrap Intervalle betrachtet [Efr79]. Ein Beispiel ist in Abbildung 4.2 in der rechten Grafik für die Eigenwerte von f_{1500} dargestellt. Bei der Erstellung von Bootstrap Intervallen wird die Matrix \hat{C} mehrfach mit unterschiedlichen Punkten $\{x'_i\}_{i=1}^M \subset \mathcal{X}$, allerdings mit gleichbleibenden Wertepaaren $\{(x_j, q_j)\}_{j=1}^N$, berechnet. Die größten und kleinsten Eigenwerte von allen berechneten Matrizen zu den einzelnen Indizes bestimmen

das Bootstrap Intervall. In der obigen Abbildung 4.2 liegt auch nach Betrachtung des Bootstrap Intervalls die Wahl eines eindimensionalen aktiven Unterraums nahe. \triangleleft

An dieser Stelle tritt die Frage auf, wie sich nun feststellen lässt, ob ein eindimensionaler Unterraum das Verhalten der Funktion f_{1500} hinreichend gut beschreiben kann. Dies lässt sich anhand der Eigenwerte nicht direkt ablesen, denn sämtliche Eigenwerte der Matrix \hat{C} sind positiv, das heißt, die Funktion ist entlang der inaktiven Richtungen $\hat{w}_2, \dots, \hat{w}_{27}$ nicht konstant. Eine grafische Methode, um die mögliche Dimensionsreduktion durch den eindimensionalen aktiven Unterraum sichtbar zu machen, und damit zu beurteilen, ob dieser Unterraum zur Dimensionsreduktion geeignet ist, wird im nächsten Abschnitt vorgestellt.

4.1.3 Sufficient summary plots

Ein Beispiel eines sogenannten *sufficient summary plots* haben wir bereits in der rechten Grafik in Abbildung 2.2 kennengelernt. Hier wurde die eindimensionale Struktur der betrachteten Funktion f sichtbar, indem der Parameterraum entsprechend der durch die Eigenvektoren der Matrix C definierte Rotation transformiert wurde. Dieses Vorgehen lässt sich wie folgt auf den Fall übertragen, dass von der Funktion f lediglich einige Wertepaare bekannt sind und die Dimension des Parameterraums \mathcal{X} größer als zwei ist. Die Grundidee der *sufficient summary plots* stammt von Cook [Coo98], der Zusammenhang zur Dimensionsreduktion findet sich in [CC02]. Der Begriff „sufficient summary plot“ lässt sich sinngemäß als „Suffizienz zusammenfassendes Diagramm“ übersetzen. Der Begriff Suffizienz stammt aus der Statistik und beschreibt die Eigenschaft einer Abbildung, hochdimensionale Daten in niedrigdimensionale Daten zu überführen, ohne grundlegende statistische Informationen zu verlieren [Fis22]. Eine grafische Möglichkeit, um zu beurteilen, ob diese Eigenschaft auf die Orthogonalprojektion in den eindimensionalen aktiven Unterraum zutrifft, stellen *sufficient summary plots* dar.

Definition 4.1 (Sufficient summary plot). *Es seien $(x_i, f(x_i))$ für $i = 1, \dots, N$ Wertepaare der Funktion $f: \mathcal{X} \rightarrow \mathbb{R}$ mit $\mathcal{X} \subseteq \mathbb{R}^m$ und $N \in \mathbb{N}$. Zudem sei $w \in \mathbb{R}^m$ mit $\|w\|_2 = 1$ gegeben. Dann wird die Grafik mit den Wertepaaren aus der Menge S_w , definiert durch*

$$S_w := \left\{ (w^T x_i, f(x_i)) \right\}_{i=1}^N,$$

als eindimensionaler sufficient summary plot der Wertepaare $(x_i, f(x_i))$, $i = 1, \dots, N$, bezüglich der Richtung w bezeichnet.

Bemerkung 4.2. Die Menge S_w von 2-Tupeln definiert im allgemeinen keine Funktion. Es kann Wertepaare $(x_1, f(x_1))$ und $(x_2, f(x_2))$ mit $w^T x_1 = w^T x_2$ geben, für die $f(x_1) \neq f(x_2)$ gilt. Die folgende Grafik veranschaulicht die Funktionsweise eines *sufficient summary plots*:

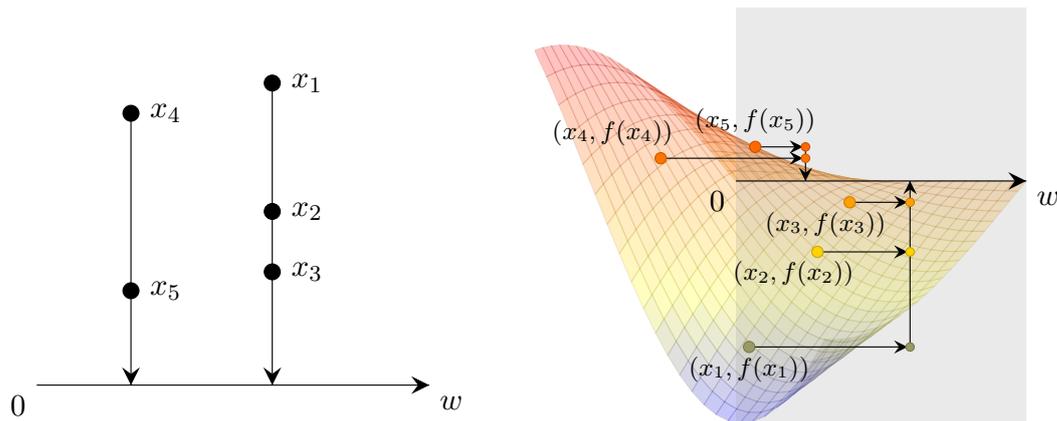


Abbildung 4.3: Schematische Illustration der Konstruktion von sufficient summary plots, die Orthogonalprojektion der Punkte auf den Vektor w (links) und die mehrdimensionale Darstellung mit zugehörigen Funktionswerten (rechts)

Die lineare Abbildung ww^T ist ein Orthogonalprojektor auf den vom Vektor w aufgespannten linearen Unterraum. Für zwei Vektoren $x_1, x_2 \in \mathbb{R}^m$ mit $ww^T x_1 = ww^T x_2$ gilt

$$\begin{aligned} ww^T x_1 &= ww^T x_2 \\ \Rightarrow w^T ww^T x_1 &= w^T ww^T x_2 \\ \Rightarrow w^T x_1 &= w^T x_2, \end{aligned}$$

denn da w normiert ist gilt $w^T w = 1$. Ist also die Orthogonalprojektion zweier Punkte auf $\text{im}(w)$ identisch, so stimmt auch die erste Koordinate der entsprechenden 2-Tupel aus S_w überein. In Abbildung 4.3 ergibt sich S_w als

$$S_w = \left\{ (c_1, f(x_1)), (c_1, f(x_2)), (c_1, f(x_3)), (c_2, f(x_4)), (c_2, f(x_5)) \right\},$$

wobei $c_1 = w^T x_1 = w^T x_2 = w^T x_3$ und $c_2 = w^T x_4 = w^T x_5$ seien. Insgesamt ergibt sich folgender sufficient summary plot:

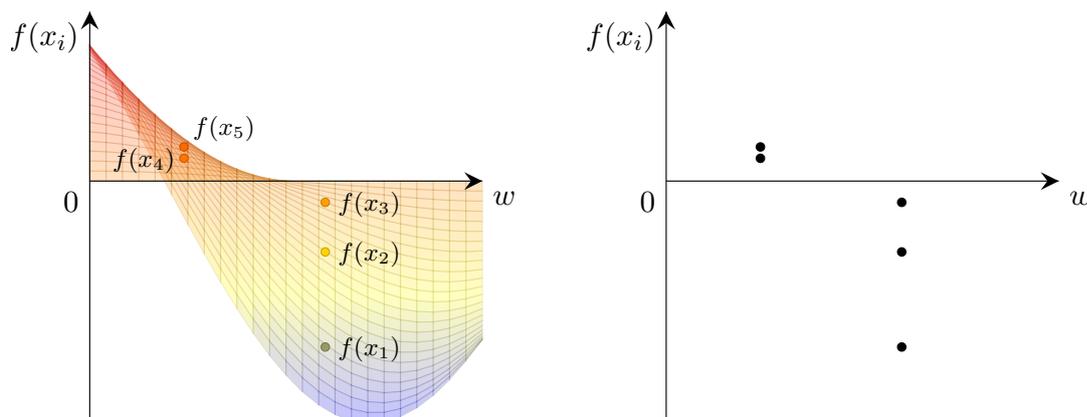


Abbildung 4.4: Punkte aus S_w zusammen mit der Funktion aus Abbildung 4.3 (links) sowie der sufficient summary plot (rechts)

◁

Wir kommen nun zurück zu dem im letzten Abschnitt berechneten eindimensionalen aktiven Unterraum \hat{w}_1 der Funktion f_{1500} und betrachten den sufficient summary plot der $N = 1000$ Wertepaare dieser Funktion bezüglich der Richtung \hat{w}_1 .

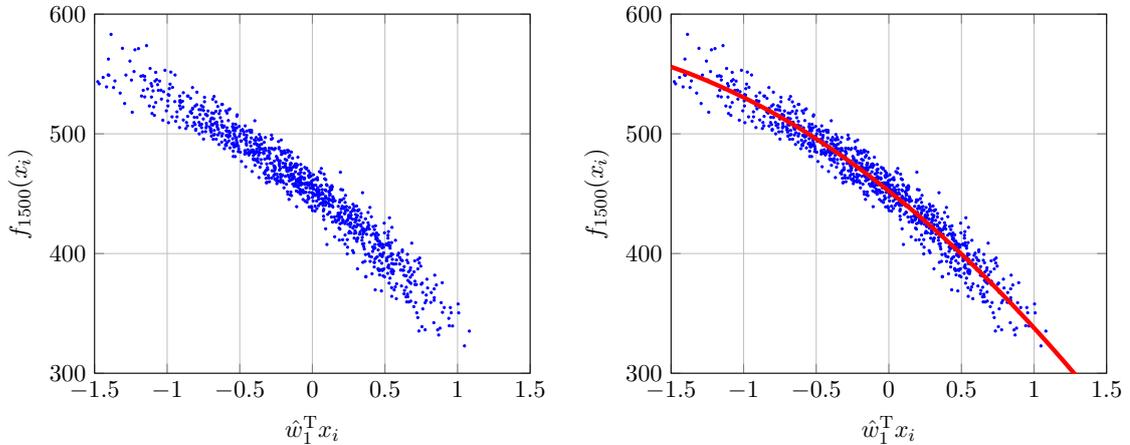


Abbildung 4.5: Sufficient summary plot von 1000 Samples der Funktion f_{1500} bezüglich der Richtung \hat{w}_1

In der obigen Abbildung stellt der sufficient summary plot, wie in Bemerkung 4.2 geschildert, keine Funktion dar, allerdings ist ein eindimensionaler Trend erkennbar. Die zu den inaktiven Richtungen gehörenden Eigenwerte sind also klein aber nicht null. Trotzdem lässt sich der sufficient summary plot der Funktion f_{1500} durch eine eindimensionale Funktion approximieren. In der rechten Grafik wurde ein quadratisches Polynom mittels Methode der kleinsten Quadrate an die vorhandenen Datenpunkte angepasst. Das quadratische Polynom $g: \mathcal{Y} \rightarrow \mathbb{R}$ mit

$$g(y) \approx -18y^2 - 96y + 452$$

beschreibt den Verlauf der Funktion f_{1500} entlang der aktiven Richtung \hat{w}_1 recht gut. Der Parameterraum $\mathcal{Y} \subseteq \mathbb{R}$ der Funktion g wird hierbei so gewählt, dass $y \cdot \hat{w}_1 \in \mathcal{X}$ für alle $y \in \mathcal{Y}$ gilt (siehe auch Abschnitt 5.1).

Die folgende Abbildung zeigt den sufficient summary plot von f_{1500} bezüglich einer zufälligen Richtung $w_r \in \mathbb{R}^{27}$. Im Vergleich zu der eindimensionalen Struktur in Abbildung 4.5 zeigt diese Grafik ein deutlich erhöhtes Rauschen, die Funktionswerte sind in vertikaler Richtung erheblich weiter gestreut. Der Vektor w_r definiert somit keinen eindimensionalen aktiven Unterraum der das Verhalten der Funktion f_{1500} gut beschreibt. Insbesondere ist die Richtung w_r nicht dazu geeignet, zur Dimensionsreduktion des Parameterraums beizutragen. Dieses Beispiel zeigt, dass sufficient summary plots eine Möglichkeit bieten, einen eindimensionalen aktiven Unterraum hinsichtlich seiner Nützlichkeit zur Dimensionsreduktion zu bewerten. In Abschnitt 4.2.3 sehen wir ein Beispiel, in dem auch der sufficient summary plot bezüglich einer mit Algorithmus 4 berechneten Richtung ein relativ starkes Rauschen aufweist.

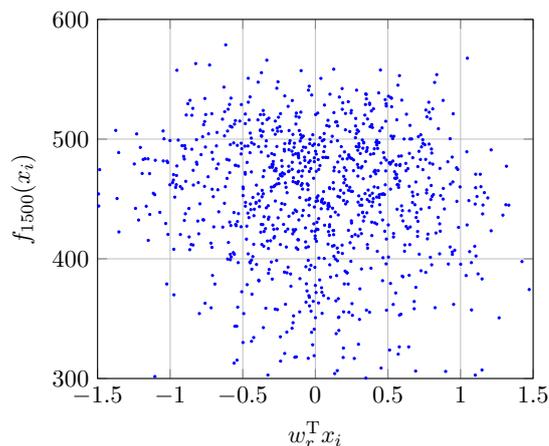


Abbildung 4.6: Sufficient summary plot von f_{1500} bezüglich der zufälligen Richtung w_r

Der eindimensionale aktive Unterraum kann in obigem Beispiel mit Hilfe eines globalen linearen Modells ebenfalls identifiziert werden. Dadurch beschränkt sich der Rechenaufwand auf die Bestimmung der Funktionswerte durch die Lösung des Differentialgleichungssystems sowie eine einzige lineare Ausgleichsrechnung. Die folgende Abbildung zeigt zum einen den sufficient summary plot bezüglich der von Algorithmus 4 produzierten Richtung \hat{w} und zum anderen die einzelnen Einträge von \hat{w} .

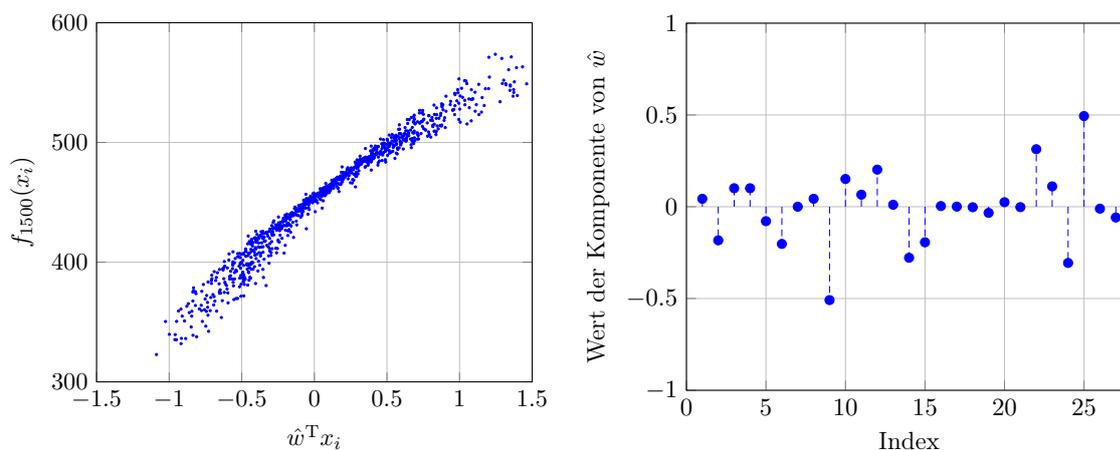


Abbildung 4.7: Sufficient summary plot bezüglich des globalen linearen Modells \hat{w} (links) und die einzelnen Einträge von \hat{w} (rechts)

Auch der sufficient summary plot bezüglich der Richtung \hat{w} zeigt eine eindimensionale Struktur. Diese ist allerdings im Vergleich zu Abbildung 4.5 entlang der vertikalen Achse durch den Ursprung gespiegelt. Würde zur Erstellung des sufficient summary plots in Abbildung 4.5 anstelle von \hat{w}_1 der Vektor $-\hat{w}_1$ genutzt, so ergäbe sich ein ähnliches Bild wie in Abbildung 4.7. Die Eigenwertzerlegung der Matrix \hat{C} ist nur bis auf das Vorzeichen der Eigenvektoren eindeutig bestimmt, sowohl \hat{w}_1 als auch $-\hat{w}_1$ sind normierte Eigenvektoren von \hat{C} , die orthogonal zu den Vektoren $\hat{w}_2, \dots, \hat{w}_m$ stehen. Daher definieren in unserem Beispiel beide Richtungen, \hat{w}_1 und $-\hat{w}_1$, denselben aktiven Unterraum. Die Orientierung von Vektoren, die einen eindimensionalen aktiven Unterraum bestimmen, wird in Abschnitt 4.1.5 noch eine zentrale Rolle spielen.

Die Komponenten des Vektors \hat{w} , die im rechten Teil von Abbildung 4.7 dargestellt sind,

geben Auskunft über den Einfluss einzelner Parameter auf den Verlauf der Funktion f_{1500} . Je kleiner der Wert \hat{w}_i eines Parameters $i \in \{1, \dots, 27\}$ im Vektor \hat{w} , desto weniger verändert sich der Funktionswert von f_{1500} in Richtung des Parameters i im Vergleich zu den orthogonalen Richtungen. Das heißt für $x \in \mathcal{X}$ ist $f_{1500}(x + \tau e_i) \approx f_{1500}(x)$, wobei e_i den i -ten Einheitsvektor bezeichnet und $\tau \in \mathbb{R}$ so gewählt ist, dass $x + \tau e_i \in \mathcal{X}$ gilt. Wie gut diese Näherung ist, hängt maßgeblich von dem zugehörigen Eigenwert λ_i sowie der Größe von τ ab, genauer gesagt von dem Produkt $\lambda_i \cdot \tau$. Anhand der Einträge von \hat{w} lassen sich also gegebenenfalls Parameter identifizieren, die nur geringfügig zum Funktionswert beitragen und die eventuell bereits in der Konstruktion des zugrundeliegenden Modells auf einen festen Wert gesetzt beziehungsweise aus dem Modell entfernt werden können. In unserem Beispiel könnten wir annehmen, dass sich die Parameter mit den Indizes 7,13,16,17,18 und 21 problemlos entfernen lassen, da diese nur sehr kleine Werte im Vektor \hat{w} aufweisen (siehe Abbildung 4.7). Dies ist allerdings nicht ohne Weiteres möglich, denn das Modell zur Simulation der T-Helferzellen-Population beruht auf allen 27 Parametern und versucht, den gesamten Krankheitsverlauf widerzuspiegeln. Zu einem anderen als dem hier betrachteten Zeitpunkt $t = 1500$, beispielsweise einige Tage nach der Infektion oder in der akuten Phase der AIDS-Erkrankung, könnten die genannten Parameter durchaus einen Einfluss auf die Zahl der T-Helferzellen haben. Das reduzierte Modell würde daher nicht mehr sämtliche Phasen des Krankheitsverlaufs hinreichend genau beschreiben. Loudon und Pankavich ([LP17], Abschnitt 3.4) konstruieren aus diesem Grund reduzierte Modelle für die einzelnen Phasen der Erkrankung, die unabhängig von den anderen Phasen sind.

Bemerkung 4.3. In [LP17] tritt zum Zeitpunkt $t = 2000$ ein zweidimensionaler aktiver Unterraum auf. In der hier gewählten Implementierung ist dies nicht der Fall. Ein Grund für diesen Unterschied könnte die OCTAVE-Funktion `ode45` zum Lösen des Differentialgleichungssystems darstellen, welche Loudon und Pankavich nicht verwenden. Auch der zweidimensionale aktive Unterraum kann mittels eines sufficient summary plots analysiert werden. Hierzu wird Definition 4.1 auf zwei Richtungen erweitert, die auf diese Weise entstehende Grafik ist somit dreidimensional. Jeder Punkt x_i wird sowohl auf die erste als auch auf die zweite aktive Richtung projiziert, x_i werden also zwei Koordinaten zugeordnet. Die dritte Koordinate ist wieder durch den Wert $f_t(x_i)$ gegeben. Ein Beispiel für einen zweidimensionalen sufficient summary plot betrachten wir in Abbildung 5.4. \triangleleft

4.1.4 Globale Approximation der T-Helferzellen-Population

Bei der Lösung des Differentialgleichungssystems bis zum Zeitpunkt $t = 1500$ wurde für gegebenes $x \in \mathcal{X}$ nicht nur $f_{1500}(x)$ approximiert, sondern auch $f_t(x)$ für $t \in \{1, \dots, 1500\}$. Zu jedem dieser Zeitpunkte können wir also einen aktiven Unterraum berechnen und den sufficient summary plot beispielsweise durch ein Polynom fünften Grades annähern. Durch dieses Vorgehen erhalten wir eine ganze Reihe von aktiven Unterräumen und Approximationen der T-Helferzellen-Population zu unterschiedlichen Zeitpunkten. Diese lassen sich anschließend zu einer globalen Approximation der T-Helferzellen-Zahl über den gesamten Zeitraum kombinieren. In Abschnitt 3.3 in [LP17] wird das genaue Vorgehen erläutert. Wir wollen an dieser Stelle nur den Ansatz und die numerischen Resultate betrachten.

Es sei $[0, T_{\text{end}}]$ mit $T_{\text{end}} \in \mathbb{N}$ das betrachtete Zeitintervall und

$$0 = t_0 < t_1 < \dots < t_{K-1} < t_K = T_{\text{end}} \quad (4.1)$$

eine Partition dieses Intervalls in $K \in \mathbb{N}$ Teilintervalle, sodass $t_i \in \mathbb{N}_0$ für alle $i = 0, \dots, K$ gilt. Zudem sei $T_i(y)$ die Approximation des sufficient summary plots zum Zeitpunkt t_i und w_i bezeichne den eindimensionalen aktiven Unterraum von f_{t_i} ermittelt durch Algorithmus 4. Hierbei wird für jedes Sample $x_j \in \mathcal{X}$, $j = 1, \dots, N$, der Wert $f_{t_i}(x_j)$ für alle $i = 0, \dots, K$ durch die approximative Lösung des Differentialgleichungssystems bis zum Zeitpunkt T_{end} berechnet. Für $t \in [0, T_{\text{end}}]$ und $x \in \mathcal{X}$ approximieren wir $T(t, x)$ durch

$$T(t, x) \approx \sum_{i=0}^K T_i(w(t)^T x) \phi_i(t), \quad (4.2)$$

wobei $w(t) = \sum_{i=0}^K w_i \phi_i(t)$ gesetzt wird. Die Hütchen-Basis $\{\phi_i\}_{i=0}^K$ auf dem durch (4.1) gegebenen Gitter ist definiert als

$$\phi_i(t) = \begin{cases} \frac{t-t_{i-1}}{t_i-t_{i-1}}, & t_{i-1} \leq t \leq t_i \\ \frac{t_{i+1}-t}{t_{i+1}-t_i}, & t_i < t \leq t_{i+1} \\ 0, & t \notin [t_{i-1}, t_{i+1}]. \end{cases}$$

Eine explizite Betrachtung der Hütchen ϕ_0 und ϕ_K an den beiden Grenzen des Intervalls wurde hier der Einfachheit halber ausgelassen.

Die Zahl der T-Helferzellen zum Zeitpunkt $t \in [0, T_{\text{end}}]$ wird in Formel (4.2) durch lineare Interpolation der approximierten Werte zu den Zeitpunkten t_l und t_{l+1} angenähert, wobei $l \in \{0, \dots, K-1\}$ so gewählt ist, dass $t_l \leq t \leq t_{l+1}$ gilt. Zudem wird zur Berechnung der aktiven Variablen $y(t) = w(t)^T x$ eine lineare Interpolation der beiden aktiven Richtungen w_l und w_{l+1} genutzt.

In der folgenden Abbildung ist sowohl die Simulation durch Lösung des Differentialgleichungssystems als auch die globale Approximation durch (4.2) dargestellt. Hierbei wurde $T_{\text{end}} = 2000$, $K = 2000$ und $t_i = i$ für $i = 0, \dots, K$ gewählt. Die Funktionen $T_i(y)$ wurden jeweils als Polynome vom Grad 5 gewählt und mit Hilfe der Methode der kleinsten Quadrate an den sufficient summary plot bezüglich w_i zum Zeitpunkt t_i angepasst. Der Parameter $x \in \mathcal{X}$ wurde uniform zufällig ausgewählt.

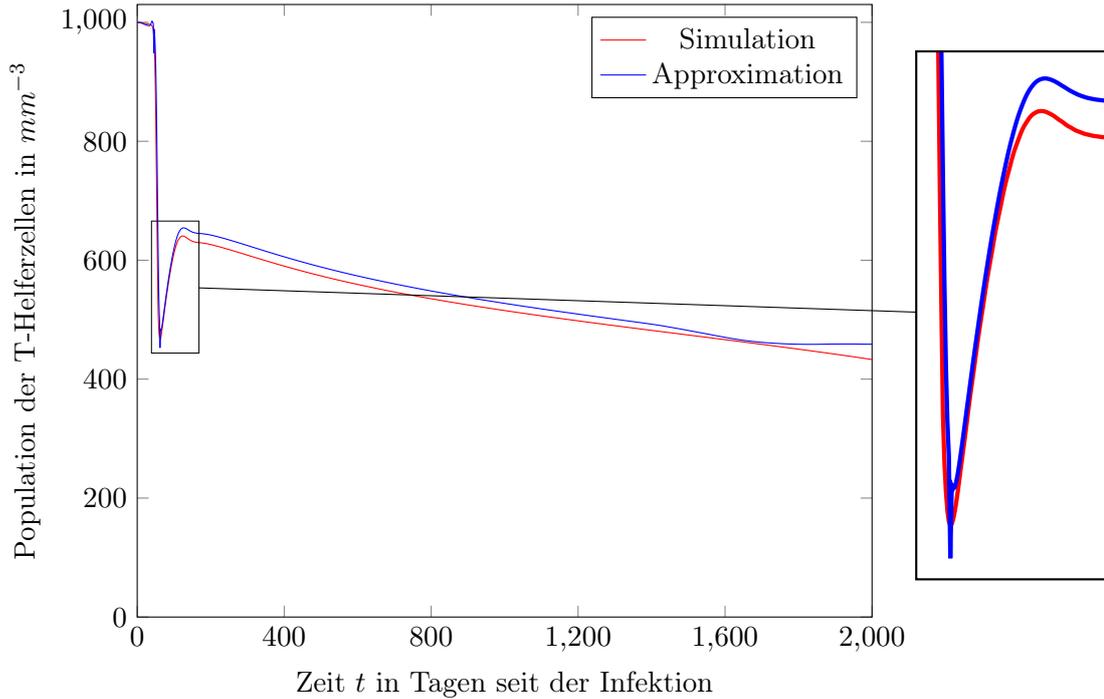


Abbildung 4.8: Globale Approximation der T-Helferzellen-Population auf dem Intervall $[0, 2000]$ sowie die Simulation aus OCTAVE

4.1.5 Kombierter aktiver Unterraum für mehrere Zeitpunkte

Wir haben uns bis dato ausschließlich darauf fokussiert, aktive Unterräume der Funktion f_t zu einem festen Zeitpunkt t zu bestimmen. In diesem Abschnitt wollen wir mehrere Zeitpunkte simultan betrachten. Das Ziel ist es, einen eindimensionalen aktiven Unterraum zu ermitteln, der das Verhalten mehrerer Funktionen f_{t_1}, \dots, f_{t_p} mit $p \in \mathbb{N}$ zu unterschiedlichen Zeitpunkten $t_1, \dots, t_p \in \mathbb{N}$ möglichst gut beschreibt. Wir beschränken uns hier auf den Fall eines eindimensionalen aktiven Unterrums, da sich die T-Helferzellen-Population zu sämtlichen Zeitpunkten recht gut durch eine eindimensionale Funktion beschreiben lässt. Damit können wir Satz 3.10 verwenden, um eine kombinierte Richtung zu bestimmen.

Wie im vorangegangenen Abschnitt seien w_1, \dots, w_p die eindimensionalen aktiven Unterräume zu den Funktionen f_{t_1}, \dots, f_{t_p} . Wir nehmen an, dass $t_i \in \{1, \dots, 2000\}$ für alle $i = 1, \dots, p$ gilt. Für diese Zeitpunkte liegen uns jeweils die Funktionswerte zu 1000 Samples vor. Die aktiven Richtungen zu den Zeitpunkten t_1, \dots, t_p werden allesamt gleich gewichtet, wir setzen also $\alpha_1 = \dots = \alpha_p = 1$.

Der kombinierte aktive Unterraum v^* wird, wie in Satz 3.10 beschrieben, als

$$v^* = \frac{\sum_{i=1}^p w_i}{\left\| \sum_{i=1}^p w_i \right\|}$$

gewählt. Um zu einem sinnvollen Ergebnis zu kommen, das heißt die sufficient summary plots zu den Zeitpunkten t_1, \dots, t_p bezüglich v^* weichen nicht zu sehr von den sufficient summary plots bezüglich w_1, \dots, w_p ab, ist die Orientierung von w_1, \dots, w_p von entschei-

dender Bedeutung. Ist beispielsweise $p = 2$ und $w_1 = -w_2$, so gilt $w_1 + w_2 = 0$. Die Aussage von Satz 3.10 ist somit nicht anwendbar. Tatsächlich kann v^* in diesem Fall als beliebiger Einheitsvektor gewählt werden, um die Summe der Abstandskvadrat zu minimieren (Satz des Thales). Allerdings definieren w_1 und w_2 denselben aktiven Unterraum, v^* sollte also ebenfalls diesen aktiven Unterraum erzeugen. Wird $w_1 = w_2$ gewählt, so verändert sich der aktive Unterraum nicht, es gilt jedoch $v^* = w_1 = w_2$ nach Satz 3.10. Die Vektoren w_1, \dots, w_p sollten somit nicht in entgegengesetzte Richtungen zeigen, damit v^* wohldefiniert ist und eine sinnvolle Näherung an die einzelnen sufficient summary plots liefert.

Die folgende Abbildung zeigt die sufficient summary plots bezüglich v^* , wobei $p = 4$ und $t_i = 1500 + 100 \cdot (i - 1)$ für $i = 1, \dots, p$ gesetzt wurde. Die Richtungen w_1, \dots, w_p wurden mit Hilfe eines globalen linearen Modells berechnet.

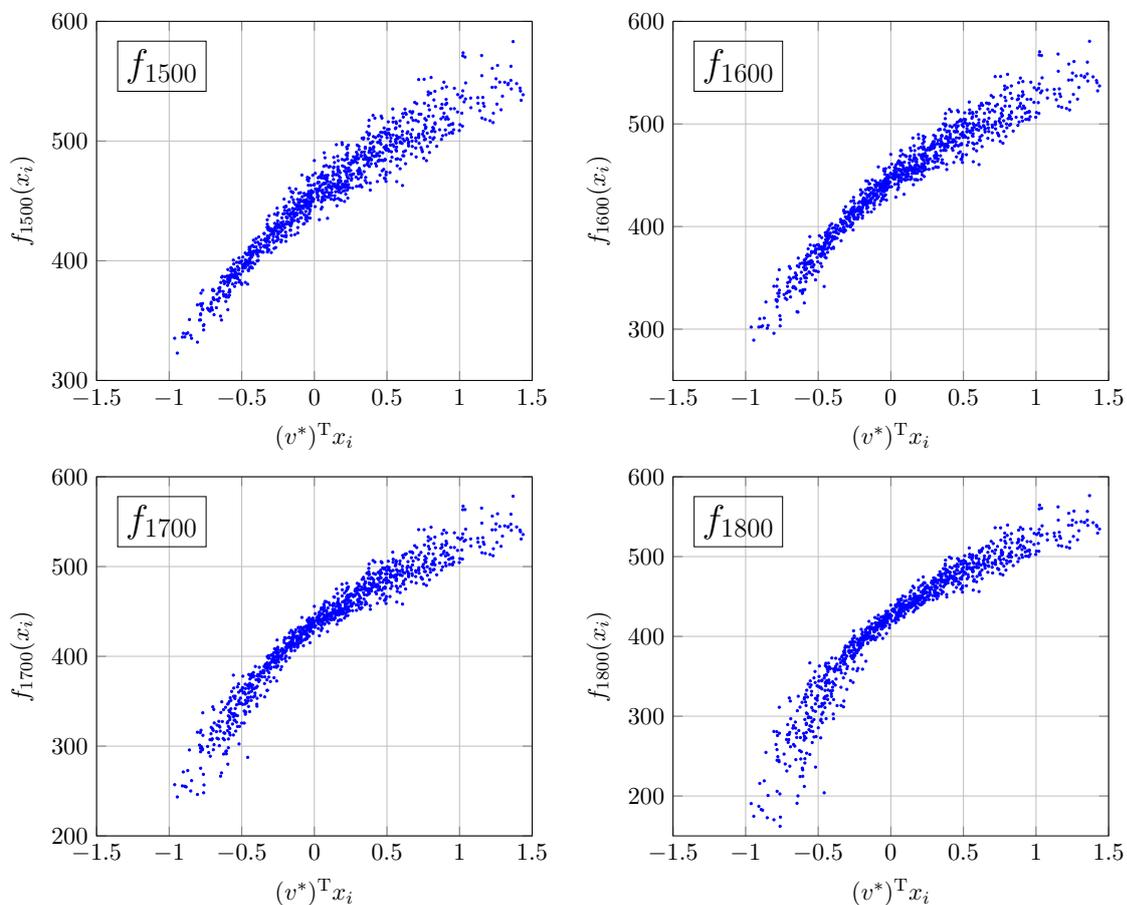


Abbildung 4.9: Sufficient summary plots von f_{1500} (oben links), f_{1600} (oben rechts), f_{1700} (unten links) und f_{1800} (unten rechts) bezüglich des kombinierten aktiven Unterrraums v^*

Die obigen Grafiken (vergleiche Abbildung 4.9) lassen allesamt eine eindimensionale Struktur erkennen. Insbesondere der sufficient summary plot der Funktion f_{1500} ist relativ ähnlich zu dem Plot in Abbildung 4.7, allerdings ist insgesamt eine größere Variabilität in den Funktionswerte ersichtlich. Die eindimensionale Struktur von f_{1500} ist in Abbildung 4.7 deutlicher erkennbar, denn die kombinierte aktive Richtung v^* definiert eine etwas andere Rotation als \hat{w} aus Abschnitt 4.1.3. Dennoch zeigt die obige Abbildung auch, dass die Rich-

ung v^* die Richtungen w_1, \dots, w_4 hinreichend gut approximiert, um den Einfluss einzelner Parameter zu mehreren Zeitpunkten anhand der Komponenten von v^* zu beurteilen.

4.2 Optimierung der Profilgeometrie von Flugzeugtragflächen

In der Aerodynamik und besonders bei der Konstruktion moderner Flugzeuge ist es unerlässlich, die Geometrie der Tragflächen mit Bedacht zu wählen. Im Bereich der sogenannten *shape optimization* wird unter anderem versucht, Auftriebs- und Reibungswerte der Flügel in einem gewissen Sinne zu optimieren (vergleiche beispielsweise [LPAC14]). Zu diesem Zweck haben Lukaczyk et al. [LPAC14] sowie Grey und Constantine [GC17] jeweils aktive Unterräume eingesetzt. Die genannten aerodynamischen Beiwerte sind dimensionslose Größen, mit deren Hilfe sich die aerodynamischen Eigenschaften der unterschiedlichen Profile vergleichen lassen (siehe [Hul07], insbesondere Abschnitt 3.3). Wir beschränken uns hier auf zweidimensionale Strömungssimulationen. Lukaczyk betrachtet in [Luk15] neben zweidimensionalen auch dreidimensionale Modelle. In diesem Abschnitt werden wir zunächst den Querschnitt von Flugzeugtragflächen modellieren und anschließend Auftriebs- und Reibungskoeffizienten in Abhängigkeit von der gewählten Geometrie berechnen. Aktive Unterräume werden uns daraufhin helfen, eine niedrigdimensionale Struktur in den betrachteten Funktionen zu identifizieren. Zum Abschluss von Abschnitt 4.2.3 wird diese Struktur genutzt, um eine in gewissem Sinne optimale Flügelgeometrie zu ermitteln, wobei allerdings auf keine der in Kapitel 5 beschriebenen Methoden zurückgegriffen wird.

4.2.1 Modellierung des Flügelquerschnitts

Wir betrachten in den folgenden Abschnitten zwei Funktionen $f_l: \mathcal{X} \rightarrow \mathbb{R}$ und $f_d: \mathcal{X} \rightarrow \mathbb{R}$, die Auftrieb (lift) beziehungsweise Reibung (drag) einer vorgegebenen Flügelgeometrie $x \in \mathcal{X}$ quantifizieren. Die Werte der beiden Funktionen spielen eine entscheidende Rolle für das aerodynamische Verhalten der Flügel [Hul07]. Um einzelne Funktionswerte von f_l oder f_d zu berechnen, sind aufwändige Simulationen notwendig. Diese werden in Abschnitt 4.2.2 näher erläutert. Zunächst wollen wir in diesem Abschnitt klären, wie die Geometrie des Flügelquerschnitts erzeugt und parametrisiert wird. Zu diesem Zweck werden wir die Wahl des Parameterraums \mathcal{X} spezifizieren. Die Definition des Parameterraums \mathcal{X} hängt eng mit der Parametrisierung der Kurven zusammen, die den Querschnitt des Flügels beschreiben.

Grundlage des Flügelquerschnitts in diesem Modell ist ein NACA0015-Profil. Die Geometrie der Profile der vierstelligen NACA-Serie wird durch vier Ziffern bestimmt, wobei die ersten beiden Ziffern die Profilwölbung sowie die Wölbungsrücklage charakterisieren und die letzten beiden Ziffern die maximale Profildicke angeben (für weitere Informationen siehe [JWP33] sowie [Mor13]). Im Fall symmetrischer NACA-Profile kann folgende Formel genutzt werden, die für $x \in [0, 1]$ die halbe Profildicke $y_\tau(x)$ angibt, wobei $\tau \in [0, 1]$ die maximale Dicke des Profils im Verhältnis zur Länge der Profilhöhe darstellt:

$$y_\tau(x) = 5\tau(0.2969\sqrt{x} - 0.126x - 0.3516x^2 + 0.2843x^3 - 0.1036x^4). \quad (4.3)$$

Hierbei wird vorausgesetzt, dass die Länge der Profilhöhe auf einen Meter normiert ist. In diesem Fall gibt τ die maximale Profildicke in Metern an. Die Formel stammt aus [JWP33], wurde allerdings insofern angepasst, dass am rechten Ende des Flügelquerschnitts obere und untere Kurve zusammentreffen, das heißt, dass $y_\tau(1) = 0$ für alle $\tau \in [0, 1]$ gilt. Wir betrachten im Folgenden ein NACA0015-Profil. Die ersten beiden Ziffern sind null, damit ist das Profil symmetrisch. Die maximale Profildicke ist gegeben als 15% in Bezug auf die Länge der Profilhöhe. Die Profilhöhe wird, wie oben gefordert, auf einen Meter normiert, daher gilt $\tau = 0.15$. Die nachfolgende Abbildung zeigt das NACA0015-Profil entsprechend der Formel (4.3).

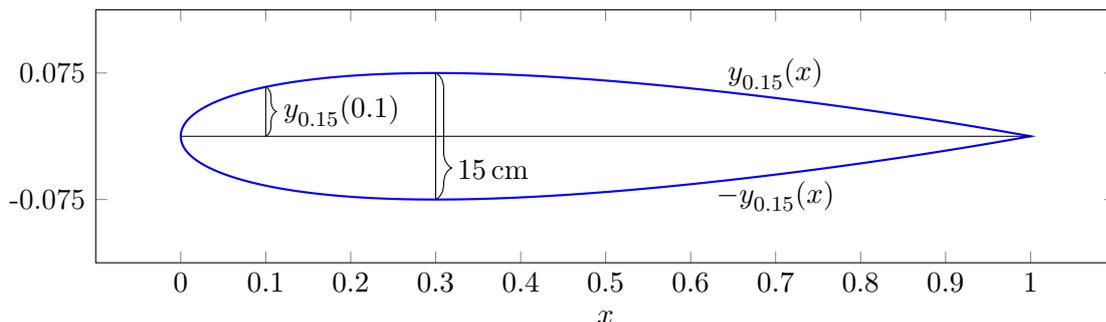


Abbildung 4.10: NACA0015-Profil

Das grundlegende NACA0015-Profil wird nun durch eine Bézierkurve verformt. Die Kurve wird sowohl zur oberen als auch zur unteren Hälfte des Profils hinzugefügt, das verformte Profil bleibt somit symmetrisch. Zunächst werden wir Bézierkurven definieren und Eigenschaften dieser Kurven betrachten, anschließend wird die genaue Konstruktion des deformierten Profils erläutert und damit der Parameterraum \mathcal{X} eingeführt.

Definition 4.2 (Bézierkurve). Die Bézierkurve $\varphi: [0, 1] \rightarrow \mathbb{R}^d$ vom Grad $m \in \mathbb{N}$ zu den Kontrollpunkten $p_0, p_1, \dots, p_m \in \mathbb{R}^d$ ist für $t \in [0, 1]$ definiert als

$$\varphi(t) := \sum_{i=0}^m B_{i,m}(t) \cdot p_i,$$

wobei

$$B_{i,m}(t) := \binom{m}{i} t^i (1-t)^{m-i}$$

das i -te Bernsteinpolynom vom Grad m bezeichnet.

In folgendem Lemma sind einige Eigenschaften von Bézierkurven zusammengetragen. Die Beweise der einzelnen Aussagen sowie weitere Bemerkungen finden sich beispielsweise in [Far14] oder [Mar06].

Lemma 4.1. Es sei $\varphi: [0, 1] \rightarrow \mathbb{R}^d$ eine Bézierkurve vom Grad $m \in \mathbb{N}$ mit den Kontrollpunkten $p_0, \dots, p_m \in \mathbb{R}^d$. Zudem sei $B_{i,m}(t)$ für $t \in [0, 1]$ das i -te Bernsteinpolynom vom Grad m . Dann gilt:

- a) Die Bézierkurve φ lässt sich rekursiv berechnen durch $\varphi(t) = b_0^m(t)$, wobei für alle

$i = 0, \dots, m$ und $k = 1, \dots, m$ gelte

$$\begin{aligned} b_i^0(t) &:= p_i, \\ b_i^k(t) &:= t \cdot b_{i+1}^{k-1} + (1-t) \cdot b_i^{k-1}. \end{aligned}$$

b) Eine Bézierkurve liegt stets in der konvexen Hülle ihrer Kontrollpunkte.

c) Bézierkurven genügen der affinen Invarianz, das heißt, für jede affine Transformation $A: \mathbb{R}^d \rightarrow \mathbb{R}^d$ und für alle $t \in [0, 1]$ gilt

$$A(\varphi(t)) = A\left(\sum_{i=0}^m B_{i,m}(t) \cdot p_i\right) = \sum_{i=0}^m B_{i,m}(t) \cdot A(p_i).$$

d) Es gilt $\varphi(0) = p_0$ und $\varphi(1) = p_m$.

Bemerkung 4.4. Die Rekursionsformel aus Lemma 4.1 a) kann genutzt werden, um auf numerisch stabile Weise für $t \in [0, 1]$ den Funktionswert $\varphi(t)$ zu berechnen. Dieses rekursive Vorgehen ist als Algorithmus von de Casteljau bekannt (siehe [Mar06], Abschnitt 6.8). Die konvexe Hülle der Kontrollpunkte wird auch als Kontrollpolygon bezeichnet. \triangleleft

Insgesamt bestimmen elf Kontrollpunkte $p_0, \dots, p_{10} \in \mathbb{R}^2$ den Verlauf der Bézierkurve $\varphi: [0, 1] \rightarrow \mathbb{R}^2$, die den Querschnitt der Tragflächen, gegeben durch das NACA0015-Profil aus Abbildung 4.10, deformiert. Anfangs- und Endpunkte werden auf die Werte

$$p_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{und} \quad p_{10} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

festgesetzt, damit obere und untere Hälfte des Flügelquerschnitts weiterhin verbunden sind. Punkt d) in Lemma 4.1 sichert, dass die Bézierkurve tatsächlich durch ihren Anfangs- und Endpunkt verläuft. Die erste Koordinate der Kontrollpunkte wird als äquidistant auf $[0, 1]$ vorgegeben. Die zweite Koordinate der Kontrollpunkte p_1, \dots, p_9 stellt jeweils einen Parameter dar. Der Parameterraum $\mathcal{X} \subseteq \mathbb{R}^9$ wird als Quader gewählt, genauer gesagt ist $x \in \mathcal{X}$ genau dann, wenn $0 \leq x_1, x_9 \leq 0.03$ und $0 \leq x_i \leq 0.07$ für alle $i = 2, \dots, 8$ gilt. Eine mögliche Bézierkurve φ zeigt Abbildung 4.11, anschließend ist in Abbildung 4.12 das durch φ deformierte Profil dargestellt, wobei dieser Flügelquerschnitt bereits um einen Winkel von $0.1396 \text{ rad} \approx 8^\circ$ gedreht wurde. Der Angriffswinkel bei der Berechnung von Reibungs- und Auftriebskoeffizienten beträgt somit 8° .

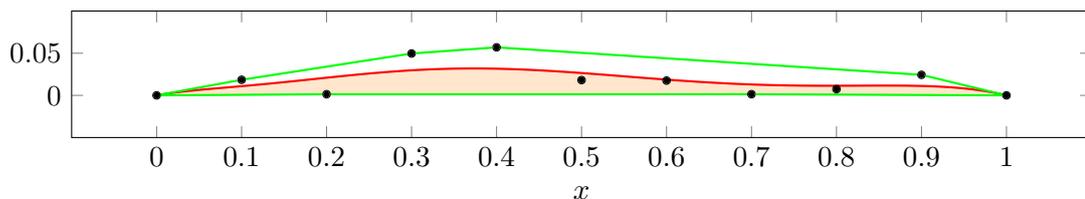


Abbildung 4.11: Beispielhafte Bézierkurve (rot) zur Deformation des NACA0015-Profiles mit den zugehörigen Kontrollpunkten (schwarze Punkte) und dem dadurch definierten Kontrollpolygon (grün)

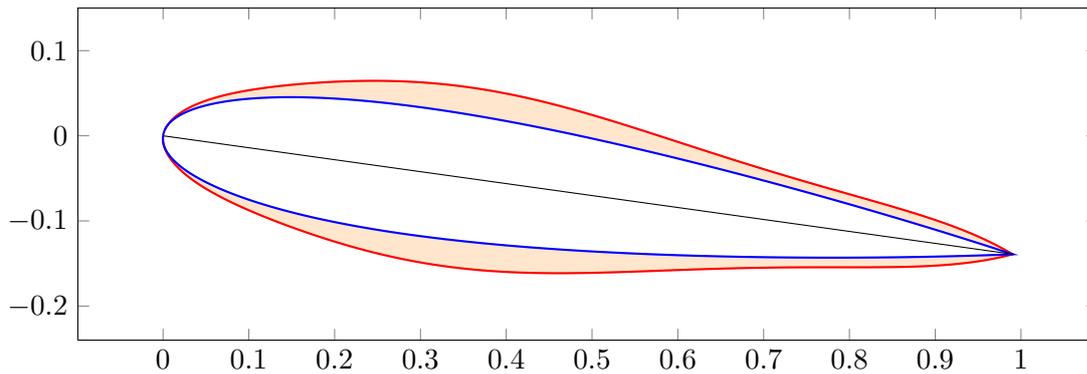


Abbildung 4.12: Rotiertes NACA0015-Profil (blau) sowie das durch die Kurve aus Abbildung 4.11 deformierte Profil (rot)

Das verformte Flügelprofil entsteht also, indem die Kurve φ aus Abbildung 4.11 zur oberen Profilhälfte, gegeben durch $y_{0.15}$ aus (4.3), addiert wird, beziehungsweise von $-y_{0.15}(x)$ subtrahiert wird, um die untere Profilhälfte zu erhalten. Wie in Abbildung 4.11 zu erkennen ist, verläuft die Bézierkurve innerhalb ihres Kontrollpolygons (siehe Lemma 4.1 b)). Dies stellt sicher, dass die Deformation durch φ zu einem zulässigen Flügelquerschnitt führt und die Deformation nicht zu stark ist. Eine negative Deformation ist in dem hier gewählten Modell nicht möglich. Die Dicke des Profils wird niemals verringert. Punkt c) in Lemma 4.1 kann genutzt werden, um zunächst die Kontrollpunkte zu rotieren und anschließend die entsprechende Kurve zu berechnen. Tatsächlich wird hier jedoch die deformierte Kurve rotiert, denn die mittels (4.3) bestimmte Funktion y_t muss ebenfalls rotiert werden. Der nächste Abschnitt zeigt, wie sich die Strömung um den deformierten Flügel sowie Auftriebs- und Reibungskoeffizienten berechnen lassen.

Bemerkung 4.5. Die Modellierung des Flügelquerschnitts und damit die Definition des Parameterraums \mathcal{X} hat großen Einfluss auf eventuell vorhandene aktive Unterräume. Es ist durchaus möglich, dass die Funktionen f_t und f_d in einem bestimmten Modell keinen aktiven Unterraum besitzen, der sich klar von den inaktiven Richtungen separieren lässt. Tatsächlich werden in [LPAC14] und [Luk15] ähnliche Kurven zur Beschreibung der Geometrie genutzt. Die exakte Modellierung ist allerdings ein wenig anders. Die hier gewählte Konstruktion ließe sich eventuell verbessern, indem die Kurve zur Deformation des NACA0015-Profiles weiter unterteilt wird. Bei Bézierkurven besteht allgemein das Problem, dass sich die Position eines einzelnen Kontrollpunktes global auf den Verlauf der erzeugten Kurve auswirkt und entsprechend auch auf die Strömungssimulation um den modellierten Flügel. \triangleleft

4.2.2 Simulation der Strömung um den Flügel

Zur Berechnung von Auftriebs- und Reibungskoeffizienten wurde die Software OPENFOAM[®] [HJF⁺12] eingesetzt ¹. Ein OCTAVE-Skript erstellt aus 1000 Punkten auf den Bézierkurven sowie 1000 Punkten des Ausgangsprofils ein Skript, welches durch Aufrufen des

¹Das OPENFOAM[®]-Tutorial zur Berechnung von Kräfte-Koeffizienten (<https://www.hpc.ntnu.no/display/hpc/OpenFOAM+-+Airfoil+Calculations>, aufgerufen am 28.06.2018) wurde mit Unterstützung von Christian Gscheidle an die hier benötigten Gegebenheiten angepasst.

OPENFOAM[®]-Befehls `snappyHexMesh` das schlussendliche Gitter erzeugt. Ein solches Gitter ist in Abbildung 4.13 (links) dargestellt. Anschließend wird mit dem Befehl `simpleFoam` die Strömungssimulation durchgeführt. Die Ermittlung der Auftriebs- und Reibungskoeffizienten kann für ein $x \in \mathcal{X}$ aus C++ heraus automatisch gestartet werden. Mehrere Shell-Skripts übernehmen den Aufruf der einzelnen Funktionen und weitere Datei-Operationen, um den OPENFOAM[®]-Case vor der Simulation zurückzusetzen. Die Koeffizienten werden mittels C++ aus einer durch OPENFOAM[®] erstellten Datei ausgelesen und zusammen mit den gewählten Parametern, die zu diesen Koeffizienten geführt haben, gespeichert. Eine einzelne Simulation dauert auf einem Kern eines Intel[®] Core i5-2430M Prozessors mit 2,4 GHz Grundtakt ungefähr 100 Sekunden. Nachfolgend ist neben dem Gitter auch die Druckverteilung um den Flügel aus Abbildung 4.12 dargestellt.

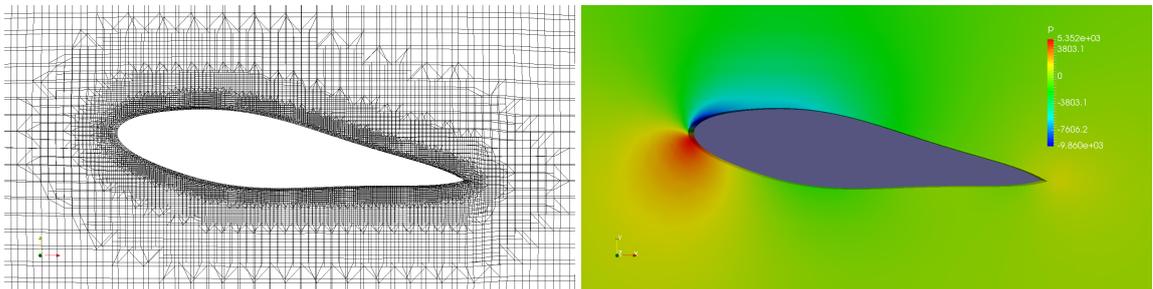


Abbildung 4.13: Gitter um den Flügel aus Abbildung 4.12 (links) und die Druckverteilung um diesen Flügelquerschnitt (rechts), Abbildung erstellt mit OPENFOAM[®]

4.2.3 Aktive Unterräume in Auftriebs- und Reibungskoeffizienten

Unser Ziel ist es nun, eine niedrigdimensionale Struktur in den Funktionen f_l und f_d , die Auftriebs- und Reibungskoeffizienten der gewählten Profile angeben, zu identifizieren. Es liegen keinerlei Informationen über die Gradienten der beiden Funktionen vor und eine Approximation mittels finiter Differenzen ist auf Grund des hohen Aufwands zur Berechnung von Funktionswerten nicht möglich. Daher greifen wir hier wieder auf lokale sowie globale Modelle für den Gradienten zurück. Die Wahrscheinlichkeitsdichte ρ sei eine Gleichverteilung auf \mathcal{X} . Zunächst wurde die Dimension der aktiven Unterräume zu beiden Funktionen mit Hilfe lokaler linearer Modelle bestimmt. Insgesamt standen zu diesem Zweck die Funktionswerte beider Funktionen zu $N = 300$ Samples zur Verfügung. Bei der Konstruktion der lokalen Modelle wurde $M = 100$ und $p = 40$ gesetzt. Wir bezeichnen mit \hat{C}_l beziehungsweise \hat{C}_d die mit Algorithmus 3 berechneten Matrizen zu den Funktionen f_l beziehungsweise f_d . Die eindimensionalen aktiven Unterräume $\hat{w}_l \in \mathbb{R}^9$ sowie $\hat{w}_d \in \mathbb{R}^9$ wurden mittels eines globalen Modells, wie in Algorithmus 4 beschrieben, berechnet. Die folgenden Abbildungen zeigen die Eigenwerte von \hat{C}_l und \hat{C}_d sowie die sufficient summary plots bezüglich der Richtungen \hat{w}_l und \hat{w}_d .

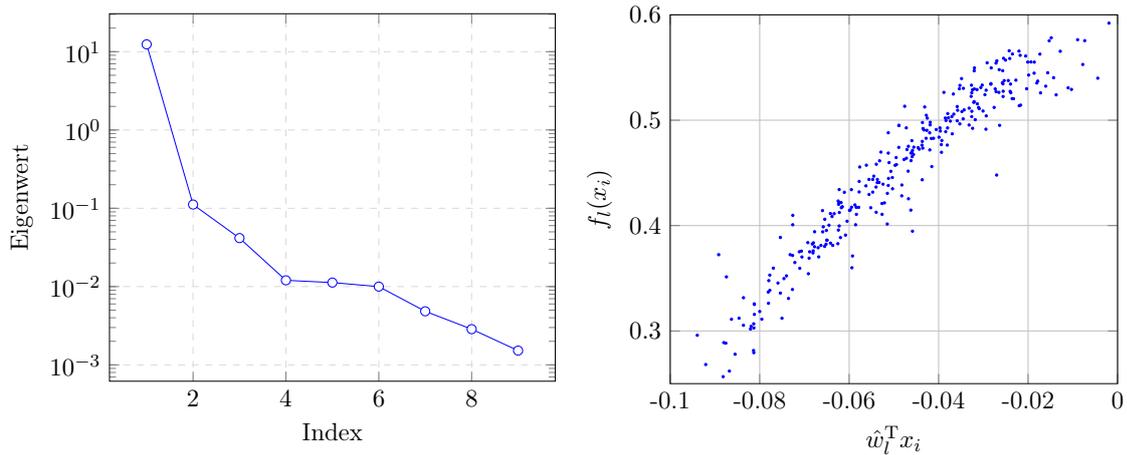


Abbildung 4.14: Eigenwerte von \hat{C}_l (links) und der sufficient summary plot der Auftriebskoeffizienten bezüglich des eindimensionalen aktiven Unterraums \hat{w}_l (rechts)

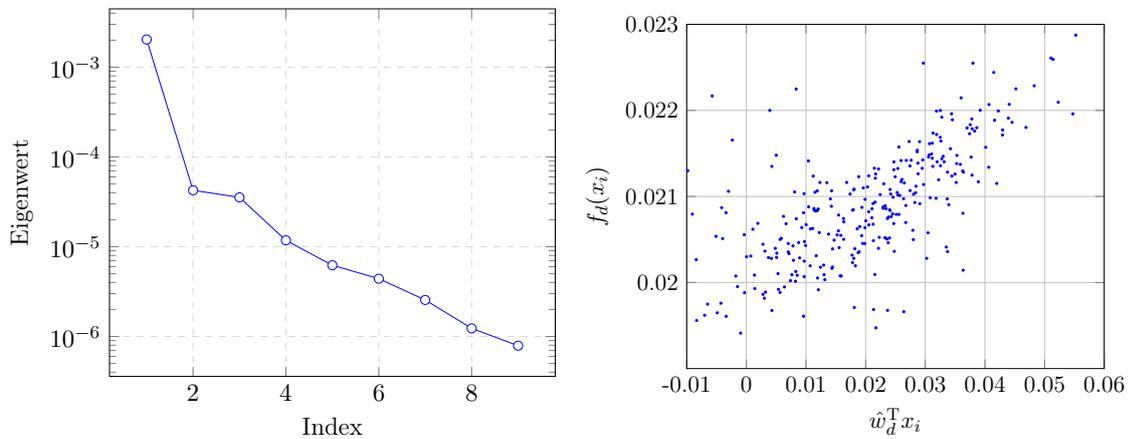


Abbildung 4.15: Eigenwerte von \hat{C}_d (links) und der sufficient summary plot der Reibungskoeffizienten bezüglich des eindimensionalen aktiven Unterraums \hat{w}_d (rechts)

Die Eigenwerte legen in beiden Fällen, das heißt sowohl für die Auftriebsbeiwerte f_l als auch für die Reibungskoeffizienten f_d , einen eindimensionalen aktiven Unterraum nahe. Die sufficient summary plots bestätigen diese Annahme, allerdings ist die Verbreitung der Funktionswerte um die eindimensionale Struktur in den Reibungskoeffizienten (Abbildung 4.15) relativ stark. Mögliche Gründe hierfür könnten Ungenauigkeiten in den berechneten Reibungsbeiwerten sein, allerdings ist auch denkbar, dass die Variabilität der Funktionswerte entlang der inaktiven Richtungen der Funktion f_d stärkere Auswirkungen hat, als beispielsweise bei den Auftriebsbeiwerten in Abbildung 4.14.

Wir betrachten nun die obigen Abbildungen im Hinblick auf eine optimale Flügelform. Ein Flügelprofil ist in diesem Zusammenhang optimal, wenn es den Auftriebskoeffizienten maximiert und gleichzeitig den Reibungskoeffizienten minimiert. Im Allgemeinen stellt diese Formulierung kein wohldefiniertes Optimierungsproblem dar, denn ein Profil mit hohem Auftriebskoeffizienten könnte einen hohen Reibungskoeffizienten aufweisen und anders herum, das heißt, die Optimierung der Funktionen f_l und f_d führt zu unterschiedlichen Flügelgeometrien. In unserem speziellen Modell ist dies nicht der Fall. Der nicht deformierte

Flügel, also das ursprüngliche NACA0015-Profil, minimiert den Reibungsbeiwert und maximiert den Auftriebskoeffizienten unter den betrachteten Flügelgeometrien. Der Verlauf der sufficient summary plots legt allerdings nahe, dass eine negative Deformation, wie in Abschnitt 4.2.1 bereits angedeutet, Auftriebs- und Reibungskoeffizienten weiter optimiert. Eine solche veränderte Modellierung könnte jedoch das soeben beschriebenen Problem, dass das Optimierungsproblem nicht wohldefiniert ist, hervorrufen.

Die identifizierten Strukturen in den Auftriebs- und Reibungskoeffizienten ermöglichen eine Reduktion des Parameterraums; die genaue Verwendung aktiver Unterräume in Parameterstudien ist Thema des nächsten Kapitels.

Anwendung auf Optimierungsprobleme

Im letzten Kapitel haben wir gesehen, wie sich aktive Unterräume in Anwendungen berechnen lassen. Allerdings wurde noch nicht erläutert, wie aktive Richtungen eingesetzt werden können, um eine Dimensionsreduktion in Parameterstudien durchzuführen. Die meisten Verfahren für Parameterstudien wie Inversion, Optimierung, Integration oder Regression benötigen eine Vielzahl an Funktionsauswertungen, wenn die Dimension m des Parameter-raums \mathcal{X} groß ist. Besonders problematisch wird dieser Umstand, wenn die Berechnung von einzelnen Funktionswerten bereits äußerst aufwändig ist. Wir möchten daher aktive Unterräume nutzen, um ansonsten unmögliche Parameterstudien durchführen zu können. Das folgende Kapitel diskutiert zunächst grundlegende Ansätze zur Verwendung aktiver Unterräume in Parameterstudien und legt anschließend einen Fokus auf Optimierungsprobleme. Im Laufe des Kapitels werden wir auf zahlreiche Probleme stoßen, die bei der Nutzung aktiver Unterräume auftreten können und werden erste Lösungsansätze kennenlernen. Viele der noch offenen Fragen sind Thema der aktuellen Forschung, in der unter anderem nach fundierten theoretischen Aussagen gesucht wird, die das gute Verhalten aktiver Unterräume in praktischen Anwendungen mathematisch stützen. Für die Konstruktion sogenannter *response surfaces*, das heißt mit Hilfe einer gegebenen Menge von Wertepaaren $\{(x_j, f(x_j))\}$ wird der Wert $f(x)$ für $x \notin \{x_j\}$ approximiert, liegen einige solcher Aussagen vor (siehe [Con15]). Regression und Interpolation fallen genau in diese Kategorie. Die anderen oben genannten Parameterstudien sind diesbezüglich weniger umfassend untersucht.

5.1 Aktive Unterräume in Parameterstudien

Zunächst wollen wir noch einmal zusammenfassen, was uns zu diesem Zeitpunkt gegeben ist. Ausgehend von einer Funktion $f: \mathcal{X} \rightarrow \mathbb{R}$ mit $\mathcal{X} \subseteq \mathbb{R}^m$ haben wir aktive Richtungen bestimmt, entlang derer sich f stärker verändert als entlang der orthogonalen Richtungen. Die Matrix C sowie die aktiven Richtungen wurden entweder mittels Algorithmus 2, falls Gradienten vorhanden sind, oder mit Algorithmus 3 beziehungsweise Algorithmus 4 und einem linearen Modell für den Gradienten bestimmt. Im Folgenden gehen wir davon aus, dass die Spalten von $W_1 \in \mathbb{R}^{m \times n}$ die aktiven Richtungen und die Spalten von $W_2 \in \mathbb{R}^{m \times (m-n)}$ die inaktiven Richtungen darstellen, die Matrix $W = \begin{bmatrix} W_1 & W_2 \end{bmatrix}$ ist die Eigenvektormatrix

von C . Die zugehörigen Eigenwerte der Matrix C seien $\lambda_1 \geq \dots \geq \lambda_m \geq 0$.

In Abschnitt 2.2.2 haben wir gesehen, dass wir jedes $x \in \mathcal{X}$ als

$$x = W_1 y + W_2 z \tag{5.1}$$

schreiben können mit $y = W_1^T x \in \mathbb{R}^n$ und $z = W_2^T x \in \mathbb{R}^{m-n}$. Damit gilt insbesondere $f(x) = f(W_1 y + W_2 z)$.

Bemerkung 5.1. Unabhängig von der betrachteten Parameterstudie können die Wertepaare, die zur Bestimmung der aktiven Unterräume berechnet wurden, für sämtliche Parameterstudien genutzt werden. Dies verringert nach Möglichkeit die Zahl der zusätzlich durchzuführenden Simulationen. \triangleleft

Ein naheliegender Ansatz wäre es, die Funktion f durch eine Funktion $g: \mathcal{Y} \rightarrow \mathbb{R}$ zu approximieren, die auf einer Menge $\mathcal{Y} \subseteq \mathbb{R}^n$ definiert ist. Die Dimension des Parameterraums der betrachteten Funktion wäre somit von m auf n reduziert worden, ganz unabhängig von der durchzuführenden Parameterstudie. Damit würde sich auch der Aufwand zur Durchführung der einzelnen Parameterstudien reduzieren. Die Konstruktion der Funktion g sollte naheliegender Weise insbesondere die aktiven Richtungen von f beachten, entlang dieser verändert sich die Funktion f im Schnitt am stärksten.

Ein erster Ansatz zur Definition der Funktion g wäre, die inaktiven Richtungen komplett zu ignorieren. Wir könnten also $z = 0$ fordern und jedes $x \in \mathcal{X}$ auf seine aktiven Variablen $y \in \mathbb{R}^n$ beschränken. Die Veränderungen von f in der inaktiven Richtung, definiert durch W_2 , würden sich somit nicht in g wiederfinden. Für einige Parameterstudien mag dieser Ansatz zu beschränkend sein, die Ergebnisse könnten sich je nach Verhalten von f entlang der inaktiven Richtungen als äußerst unbefriedigend herausstellen. Trotzdem möchten wir auf diesen Ansatz an dieser Stelle etwas genauer eingehen, hier zeigt sich nämlich bereits ein erstes Problem, das sich auch in anderen Konstruktionen wiederfindet. Es sei also g als

$$g(y) := f(W_1 y)$$

gegeben. Da $z = 0$ ist, erhalten wir aus (5.1) die Darstellung $x = W_1 y$, also entsprechend $y = W_1^T x$ aufgrund der Spalten-Orthogonalität von W_1 . Wir approximieren nun $f(x)$ durch

$$f(x) \approx g(y) = g(W_1^T x) = f(W_1 W_1^T x). \tag{5.2}$$

Hier tritt wieder der Orthogonalprojektor $W_1 W_1^T$ auf den aktiven Unterraum, definiert durch W_1 , auf (siehe Bemerkung 2.2). Problematisch wird dieser Ansatz vor allem dann, wenn $\text{im}(W_1) \not\subseteq \mathcal{X}$ gilt, das heißt wenn die Orthogonalprojektion $W_1 W_1^T x$ von $x \in \mathcal{X}$ auf $\text{im}(W_1)$ nicht unbedingt selbst in \mathcal{X} liegt. In diesem Fall wäre $f(W_1 W_1^T x)$ nicht bekannt und daher auch $g(y) = g(W_1^T x)$ nicht wohldefiniert. Folgende Abbildung veranschaulicht dieses Problem anhand der Funktion $f(x_1, x_2) = \cos(2x_1 + 5x_2)$, die bereits in Abschnitt 2.2.3 betrachtet wurde. Der Definitionsbereich von f ist $\mathcal{X} = [-1, 1]^2$ und der eindimensionale aktive Unterraum von f berechnet sich als $w_1 = \frac{1}{\sqrt{29}} [2 \ 5]^T$. Des Weiteren ist der

Orthogonalprojektor auf das Bild von w_1 durch

$$w_1 w_1^T = \frac{1}{29} \begin{bmatrix} 4 & 10 \\ 10 & 25 \end{bmatrix}$$

gegeben. Für den Punkt $x = [0.9 \ 0.9]^T$ ergibt sich als Projektion der Punkt

$$w_1 w_1^T x = \frac{1}{29} \begin{bmatrix} 4 & 10 \\ 10 & 25 \end{bmatrix} \begin{bmatrix} 0.9 \\ 0.9 \end{bmatrix} = \begin{bmatrix} 63/145 \\ 63/58 \end{bmatrix} \notin \mathcal{X}.$$

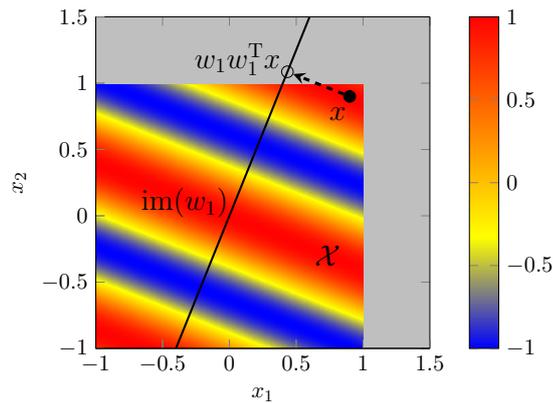


Abbildung 5.1: Beispielhafte Darstellung des Problems, dass die Orthogonalprojektion eines Punktes $x \in \mathcal{X}$ selbst nicht in \mathcal{X} liegt

Der Parameterbereich \mathcal{Y} von g sollte also so gewählt sein, dass $W_1 y \in \mathcal{X}$ für alle $y \in \mathcal{Y}$ gilt. Damit ergibt sich $\mathcal{Y} \subseteq \mathbb{R}^n$ als

$$\mathcal{Y} := \{y \in \mathbb{R}^n : W_1 y \in \mathcal{X}\}. \quad (5.3)$$

Die Parameterstudien würden nun auf der Funktion g und damit auf dem Parameterraum \mathcal{Y} mit Dimension n durchgeführt. Um tatsächlich den Funktionswert $g(y)$ für ein $y \in \mathcal{Y}$ berechnen zu können, muss y zunächst in den Parameterraum \mathcal{X} von f abgebildet werden. Hierzu wird wie oben beschrieben die Abbildung $y \mapsto W_1 y$ verwendet. Zusammenfassend ergibt sich g als

$$\begin{aligned} g: \mathcal{Y} &\rightarrow \mathbb{R} && \text{mit } \mathcal{Y} = \{y \in \mathbb{R}^n : W_1 y \in \mathcal{X}\}, \\ g(y) &= f(W_1 y). \end{aligned} \quad (5.4)$$

Die Funktion g kann für Parameterstudien genutzt werden, allerdings lässt sie keine Approximation von f auf ganz \mathcal{X} zu. Wenn wir $f(x) \approx g(W_1^T x)$ wie in (5.2) setzen, so ist dies nicht zwangsläufig für alle $x \in \mathcal{X}$ wohldefiniert. Zudem können wir nicht annehmen, dass $\lambda_{n+1} = \dots = \lambda_m = 0$ gilt. Ist dies der Fall, so gilt $f(x + W_2 z) = f(x)$ für alle $x \in \mathcal{X}$ und $z \in \mathbb{R}^{m-n}$ mit $x + W_2 z \in \mathcal{X}$. Im Allgemeinen sind die Eigenwerte zum inaktiven Unterraum allerdings nicht exakt null, die Funktion f ist dementsprechend also auch nicht konstant entlang der inaktiven Richtungen. Daher ist es für die betrachtete Parameterstudie meist ungünstig, $z = 0$ festzusetzen. Das Verhalten von f entlang der inaktiven Richtungen sollte

auf geeignete Weise in die Konstruktion der Funktion g eingehen. Eine sinnvolle Definition von g hängt stark von der Parameterstudie ab, die durchgeführt werden soll.

Wir betrachten noch einmal obiges Beispiel, also die Funktion $f: [-1, 1]^2 \rightarrow \mathbb{R}$ mit $f(x_1, x_2) = \cos(2x_1 + 5x_2)5$. Der Parameterraum \mathcal{Y} ist entsprechend der Formel (5.3) gegeben durch $\mathcal{Y} = \{y \in \mathbb{R} : w_1 y \in \mathcal{X}\}$. Dem Punkt $y = w_1^T x$ mit $x = \begin{bmatrix} 0.9 & 0.9 \end{bmatrix}^T$ ist also kein Wert $g(y)$ zugeordnet. Es wäre allerdings naheliegend, den Wert $g(y)$ beispielsweise durch $f(x)$ zu approximieren, denn es gilt immerhin $w_1 y = w_1 w_1^T x$ und es wurde $z = 0$ gesetzt (vergleiche auch Abbildung 5.1). Es ist also durchaus möglich, dem Punkt $y = w_1^T x$ einen sinnvollen Wert zuzuweisen. Die Voraussetzung, dass $W_1 y \in \mathcal{X}$ für alle $y \in \mathcal{Y}$ gelten soll, ist daher eine zu starke Einschränkung. Tatsächlich genügt es, wenn wir fordern, dass zu jedem $y \in \mathcal{Y}$ ein $x \in \mathcal{X}$ mit der Eigenschaft $y = W_1^T x$ existiert. Daher werden wir den Parameterraum \mathcal{Y} der Funktion g definieren als

$$\mathcal{Y} := \{y \in \mathbb{R}^n : y = W_1^T x \text{ für ein } x \in \mathcal{X}\}. \quad (5.5)$$

Die Definition von g aus (5.4) lässt sich nun allerdings nicht auf den modifizierten Parameterraum übertragen. Daher werden wir alternative Konstruktionen der Funktion g kennenlernen.

Im restlichen Teil dieses Abschnitts werden wir einen Ansatz betrachten, der sich für Studien wie Integration, statistische Inversion oder die Erstellung von response surfaces eignet. Mit der Optimierung wird sich der anschließende Abschnitt befassen.

Constantine [Con15] schlägt vor, die Funktion g zu konstruieren, indem die Werte von g entlang der inaktiven Richtungen gemittelt werden. Die Funktion $g: \mathcal{Y} \rightarrow \mathbb{R}$ ist für ein $y \in \mathcal{Y}$ definiert als das Integral über alle $x \in \mathcal{X}$ mit $y = W_1^T x$. Hierbei wird die Definition von \mathcal{Y} aus (5.5) genutzt, diese wird auch in den folgenden Abschnitten den Parameterraum von g charakterisieren. Um g formal angeben zu können, benötigen wir bedingte Wahrscheinlichkeitsdichten, die den Zusammenhang zwischen aktiven und inaktiven Variablen widerspiegeln. Die gemeinsame Dichte $\pi(y, z)$ der aktiven Variablen y und der inaktiven Variablen z ist durch die Wahrscheinlichkeitsdichte ρ gegeben, das heißt für alle $y \in \mathbb{R}^n$ und $z \in \mathbb{R}^{m-n}$ gilt

$$\pi(y, z) = \begin{cases} \rho(W_1 y + W_2 z) & \text{für } W_1 y + W_2 z \in \mathcal{X}, \\ 0 & \text{für } W_1 y + W_2 z \notin \mathcal{X}. \end{cases} \quad (5.6)$$

Wir fassen nun Y beziehungsweise Z als Zufallsvariablen mit Werten in \mathbb{R}^n beziehungsweise \mathbb{R}^{m-n} mit den Randdichten

$$\pi_Y(y) = \int \rho(W_1 y + W_2 z) dz \quad \text{und} \quad \pi_Z(z) = \int \rho(W_1 y + W_2 z) dy$$

auf. Die bedingte Dichte $\pi_{Z|Y}(z|y)$ von Z gegeben Y ergibt sich dann für $y \in \mathbb{R}^n$ mit $\pi_Y(y) \neq 0$ und $z \in \mathbb{R}^{m-n}$ als

$$\pi_{Z|Y}(z|y) = \frac{\pi(y, z)}{\pi_Y(y)}. \quad (5.7)$$

Wir definieren die Funktion $g: \mathcal{Y} \rightarrow \mathbb{R}$ für $y \in \mathcal{Y}$ als

$$g(y) = \int f(W_1 y + W_2 z) \pi_{Z|Y}(z|y) dz. \quad (5.8)$$

Hier tritt die Frage auf, ob g auf ganz \mathcal{Y} aus (5.5) wohldefiniert ist. Tatsächlich stellt die bedingte Dichte $\pi_{Z|Y}(z|y)$ sicher, dass der Funktionswert $f(W_1 y + W_2 z)$ nur für $W_1 y + W_2 z \in \mathcal{X}$ berechnet werden muss, außerhalb von \mathcal{X} kann f durch beliebige (endliche) Werte fortgesetzt werden. Nach (5.7) und (5.6) gilt nämlich $\pi_{Z|Y}(z|y) = 0$, falls $W_1 y + W_2 z \notin \mathcal{X}$. Die Funktion f kann nun für $x \in \mathcal{X}$ durch

$$f(x) \approx g(W_1^T x) = \int f(W_1 W_1^T x + W_2 z) \cdot \pi_{Z|Y}(z|W_1^T x) dz \quad (5.9)$$

approximiert werden. Hier tritt wieder der Orthogonalprojektor $W_1 W_1^T$ auf. Die Approximation von $f(x)$ ergibt sich nach obiger Formel also durch das Integral über die inaktiven Variablen, wobei die aktiven Variablen in Form von $W_1 W_1^T x$ festgesetzt werden. In der praktischen Umsetzung des Verfahrens kommt nun wieder ein Monte-Carlo-Ansatz zum Einsatz. Für ein festes $y \in \mathcal{Y}$ werden zufällige Punkte $\{z_i\}_{i=1}^N \in \mathbb{R}^{m-n}$ bezüglich der bedingten Dichte $\pi_{Z|Y}(z|y)$ gewählt und anschließend die entsprechenden Funktionswerte $f(W_1 y + W_2 z_i)$ der Funktion f berechnet. Die Approximation \hat{g} von g ergibt sich als arithmetisches Mittel der Funktionswerte, das heißt

$$g(y) \approx \hat{g}(y) := \frac{1}{N} \sum_{i=1}^N f(W_1 y + W_2 z_i). \quad (5.10)$$

Ein wesentliches Problem bei diesem Ansatz stellt die bedingte Wahrscheinlichkeitsdichte $\pi_{Z|Y}$ und Sampling bezüglich dieser Dichte dar. Im Allgemeinen hängt $\pi_{Z|Y}$ maßgeblich von \mathcal{X} und ρ ab und ist nicht ohne zusätzlichen Aufwand berechenbar. Bereits in dem einfachen Fall $\mathcal{X} = [-1, 1]^m$ mit der Gleichverteilung ρ auf \mathcal{X} ist die bedingte Dichte äußerst komplex (siehe [Con15], Abschnitt 4.1).

Die nachfolgende Abbildung zeigt die Funktion $f(x_1, x_2) = x_2 \cdot \cos(3x_1 + x_2)$, die aktive Richtung w_1 von f sowie beispielhafte Samples $\{W_2 z_i\}$ (grüne Punkte) entlang der inaktiven Richtung.

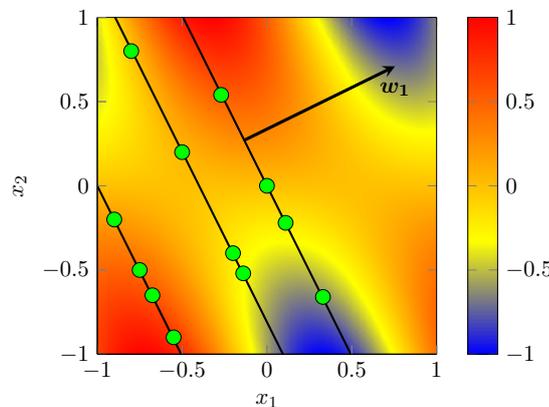


Abbildung 5.2: Beispielhafte Samples entlang der inaktiven Richtungen zur Konstruktion einer Approximation g der Funktion f

Mit (5.9) ergibt sich eine Approximation von $f(x)$ durch

$$f(x) \approx g(W_1^T x) \approx \hat{g}(W_1^T x) = \frac{1}{N} \sum_{i=1}^N f(W_1 W_1^T x + W_2 z_i).$$

Der mittlere quadratische Fehler bei der Approximation von f durch g ist beschränkt durch $C(\lambda_{n+1} + \dots + \lambda_m)^{1/2}$ mit einer Konstanten $C \in \mathbb{R}_+$ ([Con15], Theorem 4.3), der Fehler bei der Approximation von g durch \hat{g} liegt bei $\mathcal{O}(N^{-1/2})$ ([Con15], Theorem 4.4). Insgesamt ergibt sich folgende Aussage zu den Approximationseigenschaften von \hat{g} :

Satz 5.1 ([Con15], Theorem 4.4). *Für den mittleren quadratischen Fehler von \hat{g} aus (5.10) gilt*

$$\left(\int (f(x) - \hat{g}(W_1^T x))^2 \rho \, dx \right)^{1/2} \leq C_1 \left(1 + N^{-1/2} \right) (\lambda_{n+1} + \dots + \lambda_m)^{1/2},$$

wobei C_1 eine Konstante ist, die vom Parameterraum \mathcal{X} sowie der Wahrscheinlichkeitsdichte ρ abhängt.

Anhand der Fehlerabschätzung lässt sich erkennen, dass insbesondere die Eigenwerte zu den inaktiven Richtungen einen maßgeblichen Einfluss auf die Approximationsgüte von \hat{g} haben. Im Gegensatz hierzu spielen bei der Wahl der Dimension des aktiven Unterraums in Formel (2.5) die Größenordnungen der einzelnen Eigenwerte keine Rolle, einzig und allein der maximale Abstand zweier aufeinander folgender Eigenwerte ist entscheidend.

5.2 Aktive Unterräume in Optimierungsaufgaben

Optimierungsaufgaben über einem hochdimensionalen Parameterraum stellen in vielen praktischen Anwendungen eine große Schwierigkeit dar, insbesondere wenn Gradienten der betrachteten Funktion zunächst approximiert werden müssen oder aber die Berechnung von Gradienten äußerst aufwändig ist. Ohne vorherige Dimensionsreduktion des Parameterraums, beziehungsweise Approximation der zu optimierenden Funktion durch eine Funktion die von weniger Parametern abhängt, stoßen viele Verfahren in hohen Dimensionen an ihre Grenze. Wir werden in diesem Abschnitt sehen, wie die durch aktive Unterräume aufgedeckte Struktur in der Optimierung genutzt werden kann. Zur Formoptimierung von Flügelgeometrien [LPAC14],[Luk15] oder eines Schiffsrumpfes [TDG⁺18] wurden aktive Unterräume bereits eingesetzt. Die in diesem Abschnitt vorgestellten Ansätze stammen von Constantine ([Con15], Kapitel 4.4).

5.2.1 Konstruktion einer Approximation der Funktion f

Der Ansatz für die Funktion g aus Formel (5.8) ist für die Anwendung in Optimierungsaufgaben ungeeignet. Für gegebenes $y \in \mathcal{Y}$ ist der Durchschnittswert von f über den inaktiven Variablen konstant, das heißt unabhängig davon, ob es sich um ein Minimierungs- oder ein Maximierungsproblem handelt wäre der Wert $g(y)$ derselbe. Daher werden wir im Folgenden versuchen, die Funktion g so zu definieren, dass sie in der Optimierung verwendet werden kann und die Veränderungen von f entlang der inaktiven Variablen geeigneter berücksichtigt.

Wir beschränken uns im Folgenden auf Minimierungsprobleme, soll die Funktion f maximiert werden, so wird $-f(x)$ statt $f(x)$ betrachtet. Die konkrete Problemformulierung lautet somit

$$\min_{x \in \mathcal{X}} f(x). \quad (5.11)$$

Zusätzliche Nebenbedingungen an x seien vorab in die Funktion f beziehungsweise in den Definitionsbereich \mathcal{X} eingeflossen. Ein naheliegender Ansatz zur Konstruktion von g im Zusammenhang mit Optimierungsproblemen minimiert für ein festes $y \in \mathcal{Y}$ über die inaktiven Variablen, das heißt

$$g(y) = \min_{\substack{z \in \mathbb{R}^{m-n}, \\ W_1 y + W_2 z \in \mathcal{X}}} f(W_1 y + W_2 z). \quad (5.12)$$

Das Optimierungsproblem (5.11) lässt nun in das Minimierungsproblem

$$\min_{y \in \mathcal{Y}} g(y)$$

in n Variablen umschreiben. Um $g(y)$ zu berechnen, müssen wir also ein Minimierungsproblem in $m - n$ Variablen lösen. Ist n relativ klein, so ist $m - n$ recht groß und die Berechnung von $g(y)$ umso aufwändiger. Damit wird der Aufwand zur Optimierung von den ursprünglichen m Variablen x aufgeteilt in n Variablen y und weiteren $m - n$ Variablen z , allerdings ohne den Rechenaufwand zu verringern. Die obige Wahl von g führt im Allgemeinen sogar zu einem erhöhten Rechenaufwand gegenüber der direkten Optimierung von f im ganzen Parameterraum \mathcal{X} . Der Ansatz nutzt also noch nicht ausreichend die Eigenschaften von aktiven und inaktiven Unterräumen. Die Funktion f verändert sich entlang inaktiver Richtungen im Mittel weniger als entlang der aktiven Richtungen. Sind die Eigenwerte $\lambda_{n+1}, \dots, \lambda_m$ klein aber nicht null, so ist f zwar entlang von w_{n+1}, \dots, w_m nicht konstant, allerdings verändern sich die Funktionswerte auch nicht zu stark. Daher sollte in den inaktiven Richtungen und damit zur Berechnung von $g(y)$ für $y \in \mathcal{Y}$ weniger Aufwand betrieben werden, als bei der Optimierung entlang der aktiven Richtungen.

Als weiteres Problem ist nicht klar, wie die Funktion $f(W_1 y + W_2 z)$ über $z \in \mathbb{R}^{m-n}$ mit $W_1 y + W_2 z \in \mathcal{X}$ minimiert werden soll, wie also $g(y)$ berechnet werden kann. Je nachdem, welche Voraussetzungen wir an f stellen, liegen uns keine Gradienten von f und insbesondere auch keine Informationen über die zweite Ableitung von f vor. Eine Approximation von erster und zweiter Ableitung wäre sehr aufwändig. Daher kann beispielsweise das Newton-Verfahren hier nicht direkt genutzt werden.

Um den soeben beschriebenen Schwierigkeiten zu begegnen, gibt es zwei unterschiedliche Ansätze. Die erste Idee betrachtet die Konstruktion von g aus (5.12), löst allerdings das Optimierungsproblem zur Berechnung von $g(y)$ nicht exakt, sondern näherungsweise. Hierzu kommen Standard-Verfahren zum Einsatz, die beispielsweise vorzeitig abgebrochen werden. Dies beschränkt den zur Berechnung von $g(y)$ betriebenen Aufwand, die genaue Umsetzung hängt allerdings von dem eingesetzten Optimierungsverfahren ab. Der zweite Ansatz minimiert in (5.12) nicht die Funktion f über den inaktiven Variablen, sondern eine Approximation \hat{f} von f . Zur Berechnung von \hat{f} können, wie in Bemerkung 5.1 an-

gedeutet, die Wertepaare genutzt werden, die zur Bestimmung der aktiven Unterräume bereits berechnet wurden. Reichen diese Wertepaare zur Konstruktion des für \hat{f} gewählten Modells aus, so müssen keine weiteren Simulationen durchgeführt werden. Der Aufwand zur Berechnung von g beschränkt sich somit auf die Bestimmung von \hat{f} und anschließender Minimierung von \hat{f} entlang der inaktiven Richtungen. Es bleibt die Frage, wie \hat{f} konstruiert werden sollte. Constantine [Con15] wählt den Ansatz eines quadratischen Polynoms um f zu approximieren. Mittels der Methode der kleinsten Quadrate können wir Koeffizienten $A \in \mathbb{R}^{m \times m}$, $b \in \mathbb{R}^m$ und $c \in \mathbb{R}$ bestimmen, sodass

$$f(x) \approx \hat{f}(x) = \frac{1}{2}x^T A x + b^T x + c \quad (5.13)$$

gilt. Die Matrix A ist hierbei symmetrisch, daher werden mindestens $\frac{m(m+1)}{2} + m + 1 = \binom{m+2}{2}$ Wertepaare benötigt, um A , b und c zu bestimmen. Für ein mögliches Vorgehen in dem Fall, dass weniger als $\binom{m+2}{2}$ Paare vorliegen, siehe [Con15]. Für $y \in \mathcal{Y}$ ergibt sich mit der Approximation \hat{f} aus (5.13), eingesetzt in (5.12), der Wert von g als

$$\begin{aligned} g(y) &= \min_{\substack{z \in \mathbb{R}^{m-n}, \\ W_1 y + W_2 z \in \mathcal{X}}} \hat{f}(W_1 y + W_2 z) \\ &= \min_{\substack{z \in \mathbb{R}^{m-n}, \\ W_1 y + W_2 z \in \mathcal{X}}} \frac{1}{2}(W_1 y + W_2 z)^T A (W_1 y + W_2 z) + b^T (W_1 y + W_2 z) + c \\ &= \frac{1}{2}y^T W_1^T A W_1 y + b^T W_1 y + c + \min_{\substack{z \in \mathbb{R}^{m-n}, \\ W_1 y + W_2 z \in \mathcal{X}}} \frac{1}{2}z^T \tilde{A} z + \tilde{b}^T z, \end{aligned}$$

wobei $\tilde{A} = W_2^T A W_2$ und $\tilde{b} = W_2^T A W_1 y + W_2^T b$ seien. Um $g(y)$ zu berechnen, muss also ein quadratisches Polynom in $z \in \mathbb{R}^{m-n}$ minimiert werden. Welches Verfahren hierzu eingesetzt werden sollte, hängt insbesondere vom Parameterraum \mathcal{X} ab. Allerdings kann auch dieses Verfahren wieder vorzeitig abgebrochen werden, um den Rechenaufwand gering zu halten. Ist \mathcal{X} durch ein (endliches) System linearer Ungleichungs- und Gleichungsnebenbedingungen bestimmt, ist \mathcal{X} also ein Polyeder, so kann beispielsweise eine Aktive-Mengen-Strategie eingesetzt werden. Im Fall $\mathcal{X} = \mathbb{R}^m$ suchen wir für festes $y \in \mathcal{Y}$ das globale Minimum der Funktion $h: \mathbb{R}^{m-n} \rightarrow \mathbb{R}$ mit $h(z) = \frac{1}{2}z^T \tilde{A} z + \tilde{b}^T z$. Wir betrachten den Gradienten $\nabla_z h(z)$ von h , dieser ist gegeben durch

$$\nabla_z h(z) = \tilde{A} z + \tilde{b},$$

denn $\tilde{A} = W_2^T A W_2$ ist symmetrisch. Damit ist \tilde{A} die Hesse-Matrix von h und es existiert ein Minimum von h genau dann, wenn \tilde{A} positiv-semidefinit ist. Dies ist äquivalent dazu, dass A positiv-semidefinit auf dem Bild von W_2 ist. Gilt dies nicht, so ist die Funktion h nach unten unbeschränkt. Ein solches Problem kann tatsächlich auch auftreten, wenn \mathcal{X} eine echte Teilmenge von \mathbb{R}^m ist, beispielsweise ein unbeschränktes Polyeder. Die Unbeschränktheit von h schließt allerdings nicht aus, dass die Funktion f beschränkt ist, dass also ein Minimum von f existiert. Es bleiben somit noch einige Fragen zur Anwendung aktiver Unterräume in der Optimierung offen.

Bemerkung 5.2. An dieser Stelle sei noch einmal darauf hingewiesen, dass die beschriebe-

nen Ansätze zur Verwendung aktiver Unterräume in Optimierungsaufgaben in der Praxis insbesondere zum Finden von Startwerten für andere Verfahren eingesetzt werden können. Dies stellt sich in hohen Dimensionen oftmals als schwieriges Problem heraus, daher können aktive Unterräume hier gegebenenfalls einen vernünftigen Startwert liefern. \triangleleft

5.2.2 Optimierung der Flügelgeometrie unter Nebenbedingungen

Zum Abschluss von Abschnitt 4.2.3 haben wir gesehen, dass die optimale Flügelgeometrie des in Abschnitt 4.2 betrachteten Modells durch das nicht deformierte NACA0015-Profil gegeben ist, wenn wir keine weiteren Restriktionen hinzufügen und Reibungs- sowie Auftriebskoeffizienten optimieren möchten. An dieser Stelle möchten wir eine Flügelform bestimmen, die den Reibungskoeffizienten minimiert, wobei gleichzeitig der Auftriebskoeffizient des Profils durch eine Konstante beschränkt wird. Zudem soll die Nebenbedingung sowie die Zielfunktion entsprechend der Ansätze aus den vorangegangenen Abschnitten approximiert werden. Zur Bestimmung der optimalen Flügelgeometrie sollen also keine weiteren Simulationen notwendig sein.

Wir betrachten folgendes Optimierungsproblem, wobei der Auftriebskoeffizient durch 0.45 nach oben beschränkt ist:

$$\begin{aligned} \min_x \quad & f_d(x) \\ \text{sodass} \quad & x \in \mathcal{X} \\ & f_l(x) \leq 0.45 \end{aligned} \tag{5.14}$$

Um die Schwierigkeiten, die bei diesem Optimierungsproblem auftreten, besser zu erkennen, sei der Parameterraum \mathcal{X} hier noch einmal explizit angegeben,

$$\mathcal{X} = \{x \in \mathbb{R}^9 : 0 \leq x_1, x_9 \leq 0.03 \text{ und } 0 \leq x_i \leq 0.07 \text{ für alle } i = 2, \dots, 8\}.$$

Zudem bezeichnen wir mit \mathcal{X}' den Parameterraum \mathcal{X} eingeschränkt auf diejenigen Profile, die der zusätzlichen Nebenbedingung an den Auftriebskoeffizienten gehorchen, das heißt

$$\mathcal{X}' := \mathcal{X} \cap \{x \in \mathbb{R}^9 : f_l(x) \leq 0.45\}.$$

Das Optimierungsproblem aus Formel (5.14) lässt sich somit auch schreiben als

$$\min_{x \in \mathcal{X}'} f_d(x).$$

Der Quader $\mathcal{X} \subset \mathbb{R}^9$ wird somit durch eine weitere im Allgemeinen nichtlineare Nebenbedingung eingeschränkt. Aus Abbildung 4.14 geht allerdings hervor, dass sich die Nebenbedingung $f_l(x) \leq 0.45$ durchaus als lineare Nebenbedingung modellieren lässt. Wird mittels der Methode der kleinsten Quadrate eine lineare Funktion an den sufficient summary plot angepasst, so ergibt sich

$$f_l(x) \approx r(\hat{w}_l^T x), \tag{5.15}$$

wobei $r: \mathbb{R} \rightarrow \mathbb{R}$ gegeben ist als

$$r(t) = 0.63 + 3.63 \cdot t \quad (5.16)$$

und \hat{w}_l den eindimensionalen aktiven Unterraum von f_l bezeichnet.

Um zusätzliche Simulationen zu verhindern, wird die Zielfunktion f_d durch ein quadratisches Polynom $\tilde{f}_d(x) = \frac{1}{2}x^T A x + b^T x + c$ mit $A \in \mathbb{R}^{9 \times 9}$, $b \in \mathbb{R}^9$ und $c \in \mathbb{R}$ approximiert. Die Nebenbedingung $f_l(x) \leq 0.45$ wird zudem durch die Bedingung $f_l(x) \approx r(\hat{w}_l^T x) \leq 0.45$ angenähert. Der Parameterbereich von \tilde{f}_d ist somit gegeben durch

$$\mathcal{X}'_r := \mathcal{X} \cap \{x \in \mathbb{R}^9 : (3.63 \cdot \hat{w}_l)^T x \leq -0.18\}. \quad (5.17)$$

Die Ungleichung in der obigen Definition ergibt sich aus den Formeln (5.15) und (5.16), wobei die Ungleichung so umgestellt wurde, dass eine lineare Nebenbedingung in Standardform vorliegt. Insgesamt ergibt sich das folgende Minimierungsproblem:

$$\begin{aligned} \min_x \quad & \tilde{f}_d(x) = \frac{1}{2}x^T A x + b^T x + c \\ \text{sodass} \quad & x \in \mathcal{X}'_r \end{aligned} \quad (5.18)$$

Abbildung 5.3 zeigt die Gerade r und den sufficient summary plot von f_l aus Abbildung 4.14. Des Weiteren sind diejenigen Samples grau hinterlegt, die die Nebenbedingung $f_l(x) \leq 0.45$ erfüllen. Der rote Bereich kennzeichnet die Samples, die der Bedingung $r(\hat{w}_l^T x) \leq 0.45$ genügen, die ursprüngliche Nebenbedingung $f_l(x) \leq 0.45$ allerdings verletzen. Diese Samples liegen also in \mathcal{X}'_r , jedoch nicht in \mathcal{X}' . Wir werden ähnliche Probleme am Ende dieses Abschnitts kurz diskutieren.

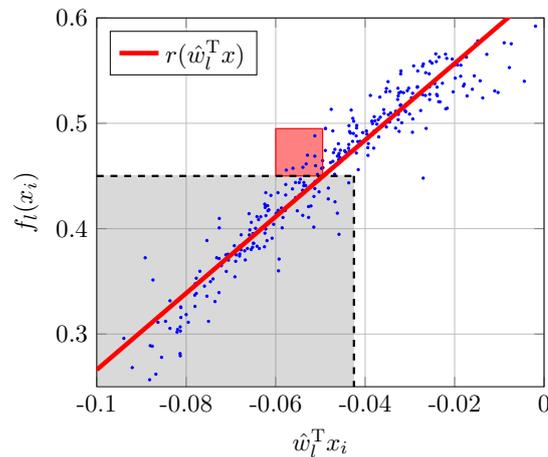


Abbildung 5.3: Sufficient summary plot von f_l (blaue Punkte), die Approximation r (rote Gerade) von f_l , der Bereich der Samples, die der tatsächlichen Nebenbedingung an den Auftriebskoeffizienten genügen (grau gefärbt) sowie der Bereich der Samples, die zusätzlich die approximierte Nebenbedingung erfüllen (rot gefärbt)

Wir betrachten in der folgenden Abbildung einen zweidimensionalen sufficient summary plot von f_l aus zwei unterschiedlichen Perspektiven. Hierbei wurden die Samples, die

bereits zur Erstellung der obigen Abbildung genutzt wurden, wiederverwendet. Die Richtungen $\hat{w}_{l,1}$ und $\hat{w}_{l,2}$ wurden mittels eines lokalen linearen Modells gemäß Algorithmus 3 berechnet. Zudem zeigt die Grafik, wie schon in Abbildung 5.3, den Bereich der Samples, die die Nebenbedingung an den Auftriebskoeffizienten erfüllen. Allerdings kann dieser Bereich nicht, wie im linken Teil der Abbildung schematisch angedeutet, durch eine Gerade von den restlichen Samples, die der Nebenbedingung nicht genügen, getrennt werden. Tatsächlich lässt der rechte Teil der Abbildung erkennen, dass sich die Funktion f_l auch entlang der inaktiven Richtung $\hat{w}_{l,2}$ verändert.

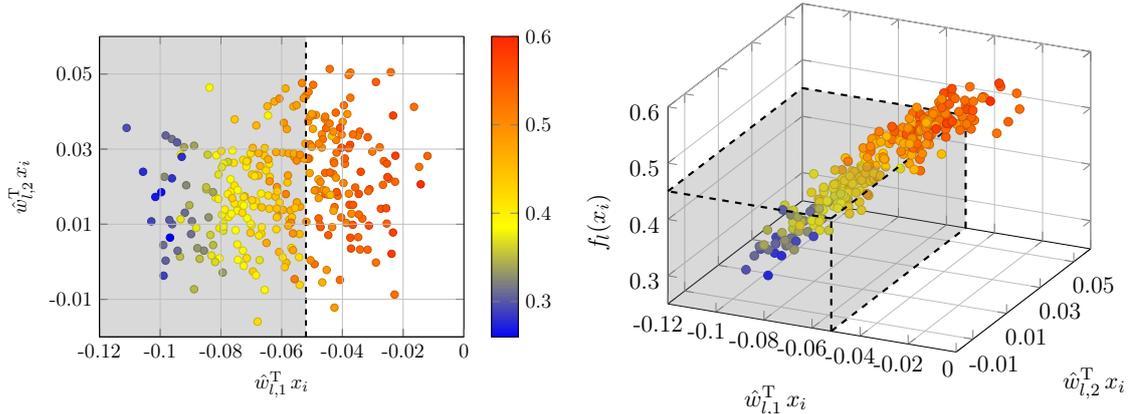


Abbildung 5.4: Zweidimensionaler sufficient summary plot von f_l bezüglich der Richtungen $\hat{w}_{l,1}$ und $\hat{w}_{l,2}$ (die unterschiedlichen Farben der Punkte geben den entsprechenden Funktionswert an) sowie der Bereich der Samples, die der Nebenbedingung an den Auftriebskoeffizienten genügen (grau gefärbt)

Die Menge \mathcal{X}'_r aus (5.17) ist durch 19 lineare Ungleichungen gegeben. Damit ist (5.18) ein quadratisches Programm in 9 Variablen. Ein solches quadratisches Optimierungsproblem kann beispielsweise durch eine Aktive Mengen Strategie gelöst werden, es wird also die Menge der in dem Optimum mit Gleichheit erfüllten Ungleichungen bestimmt (siehe [NW99], insbesondere Kapitel 16.4).

Die folgende Abbildung zeigt eine aus (5.18) berechnete optimale Flügelgeometrie, deren Auftriebskoeffizient durch 0.45 nach oben beschränkt ist und deren Reibungskoeffizient unter dieser Bedingung minimal ist. Tatsächlich liegt der mit einer weiteren Simulation berechnete Auftriebskoeffizient des Profils mit circa 0.469 leicht über 0.45, vergleiche Abbildung 5.3 (rot eingefärbter Bereich). Allerdings ist die Nebenbedingung $r(\hat{w}_l^T x) \leq 0.45$ mit Gleichheit erfüllt (siehe auch die nach der Abbildung vorgestellten Probleme der Optimierungsaufgabe (5.18)). Der Reibungskoeffizient des Profils beträgt ungefähr 0.0219 (vergleiche den sufficient summary plot in Abbildung 4.15 für Referenzwerte).

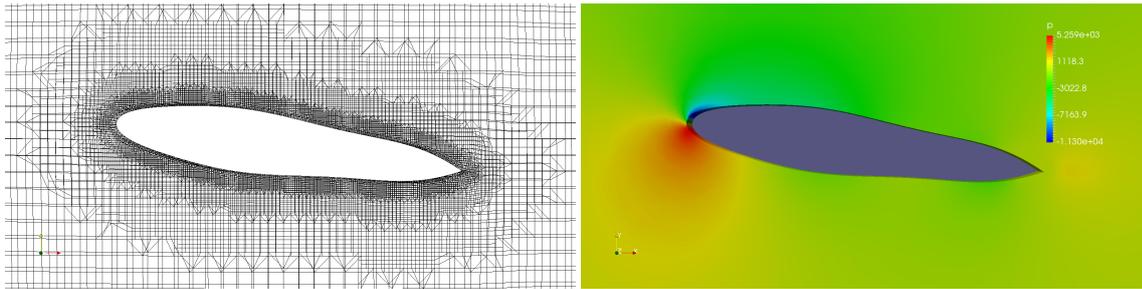


Abbildung 5.5: Gitter des optimalen Flügelprofils (links) und die Druckverteilung um diesen Flügelquerschnitt (rechts), Abbildung erstellt mit OPENFOAM®

Eine optimale Lösung x^* des Problems (5.18) kann folgende Nachteile innehaben:

- Die Zielfunktion \tilde{f}_d ist eine Approximation an die tatsächlich zu minimierende Funktion f_d , daher kann es einen Punkt $x' \in \mathcal{X}'_r$ mit $f_d(x') < f_d(x^*)$ geben, das heißt x^* ist bezüglich f_d nicht optimal.
- Der Parameterbereich \mathcal{X}'_r ist ebenfalls nur eine Approximation der Menge \mathcal{X}' . Dies kann zum einen dazu führen, dass x^* zwar in \mathcal{X}'_r optimal ist, allerdings nicht in \mathcal{X}' , das heißt es gibt $x' \in \mathcal{X}' \setminus \mathcal{X}'_r$, sodass $\tilde{f}_d(x') < \tilde{f}_d(x^*)$ gilt. Zum anderen kann der Fall eintreten, dass die berechnete Lösung x^* keine zulässige Lösung in \mathcal{X}' ist, das heißt $x^* \in \mathcal{X}'_r \setminus \mathcal{X}'$ gilt.

Gegebenenfalls sind somit weitere Untersuchungen der Lösung x^* notwendig. Insbesondere sollte geprüft werden, ob x^* eine zulässige Lösung des ursprünglichen Optimierungsproblems (5.14) ist. Sollte dies nicht der Fall sein, so könnte beispielsweise die Approximation der Nebenbedingung an den Auftriebskoeffizienten aus Formel (5.15) verbessert werden.

Zusammenfassung und Ausblick

Aktive Unterräume stellen einen vielversprechenden Ansatz zur Dimensionsreduktion in Parameterstudien dar. In der vorliegenden Arbeit wurde dieser Ansatz vorgestellt und praktisch umgesetzt. Die Implementierung wurde möglichst allgemein gehalten und gibt die Möglichkeit, sowohl Gradienten als auch lokale und globale lineare Modelle zur Approximation der Gradienten für die Berechnungen zu nutzen. Die Anwendung aktiver Unterräume auf die Simulation des Verlaufs einer HIV-Erkrankung sowie Strömungssimulationen um ein deformiertes NACA-Profil zeigt das Potential der Methode der aktiven Unterräume. Wir konnten in komplexen Modellen niedrigdimensionale Strukturen sichtbar machen und diese zur Konstruktion von Approximationen der betrachteten Funktionen nutzen. Im Bereich der Optimierung, unter Verwendung der identifizierten Strukturen, wurden erste Ansätze vorgestellt, allerdings sind in dieser Anwendung noch einige Fragen offen. Insbesondere gibt es kaum Aussagen, die das gute Verhalten der beschriebenen Verfahren in der Praxis, durch theoretische Erkenntnisse stützen.

Die Untersuchung der Ergebnisse bei der Simulation einer HIV-Erkrankung hat zudem weitere Alternativen ergeben, wie sich aktive Unterräume zur Analyse von Zeitreihen einsetzen lassen. Ein neuer Ansatz, der mehrere eindimensionale aktive Unterräume kombiniert, wurde entwickelt und implementiert sowie praktisch eingesetzt. Wir haben gesehen, dass auch die kombinierte Richtung zu ansprechenden Ergebnissen führt und konnten diese durch eine Optimalitätseigenschaft, die einen Bezug zu den ursprünglichen Richtungen herstellt, bestätigen. Weiterführende Untersuchungen in diesem Bereich könnten zu einer Verallgemeinerung des vorgestellten Verfahrens auf mehrdimensionale aktive Unterräume führen.

In der aktuellen Forschung gewinnen aktive Unterräume stetig an Bedeutung, was durch eine Vielzahl aktueller Publikationen bestätigt wird. Immer mehr Anwendungen greifen auf diese Methode zurück und immer häufiger werden aktive Unterräume erfolgreich zur Dimensionsreduktion in aufwändigen Simulationen eingesetzt. Trotzdem ist die Methode der aktiven Unterräume noch nicht vollständig untersucht und es bleiben viele interessante Ideen und Vorschläge, die eine weitere Betrachtung lohnenswert erscheinen lassen.

Abbildungsverzeichnis

2.1	Dreidimensionale Abbildung der Funktion $f(x_1, x_2) = \cos(2x_1 + 5x_2)$	11
2.2	Die Funktion $f(x_1, x_2) = \cos(2x_1 + 5x_2)$ von oben (links) und von der Seite (rechts)	12
3.1	Veranschaulichung des Abstandes zweier Unterräume (Definition 3.1)	19
3.2	Grafische Veranschaulichung der Konstruktion lokaler linearer Modelle zur Approximation von Gradienten	23
3.3	Beispielhafte Berechnung von v^* mit Satz 3.10	29
4.1	Simulationen der T-Helferzellen-Population mit unterschiedlichen Wahlen der 27 Parameter in Abhängigkeit von der Zeit t seit der Infektion	33
4.2	Eigenwerte der Matrix \hat{C} für den Zeitpunkt $t = 1500$ in absteigender Sortierung (links) und zusätzlich mit Bootstrap Intervallen (rechts)	34
4.3	Schematische Illustration der Konstruktion von sufficient summary plots, die Orthogonalprojektion der Punkte auf den Vektor w (links) und die mehrdimensionale Darstellung mit zugehörigen Funktionswerten (rechts)	36
4.4	Punkte aus S_w zusammen mit der Funktion aus Abbildung 4.3 (links) sowie der sufficient summary plot (rechts)	36
4.5	Sufficient summary plot von 1000 Samples der Funktion f_{1500} bezüglich der Richtung \hat{w}_1	37
4.6	Sufficient summary plot von f_{1500} bezüglich der zufälligen Richtung w_r	38
4.7	Sufficient summary plot bezüglich des globalen linearen Modells \hat{w} (links) und die einzelnen Einträge von \hat{w} (rechts)	38
4.8	Globale Approximation der T-Helferzellen-Population auf dem Intervall $[0, 2000]$ sowie die Simulation aus OCTAVE	41
4.9	Sufficient summary plots von f_{1500} (oben links), f_{1600} (oben rechts), f_{1700} (unten links) und f_{1800} (unten rechts) bezüglich des kombinierten aktiven Unterrums v^*	42
4.10	NACA0015-Profil	44
4.11	Beispielhafte Bézierkurve (rot) zur Deformation des NACA0015-Profiles mit den zugehörigen Kontrollpunkten (schwarze Punkte) und dem dadurch definierten Kontrollpolygon (grün)	45
4.12	Rotiertes NACA0015-Profil (blau) sowie das durch die Kurve aus Abbildung 4.11 deformierte Profil (rot)	46
4.13	Gitter um den Flügel aus Abbildung 4.12 (links) und die Druckverteilung um diesen Flügelquerschnitt (rechts), Abbildung erstellt mit OPENFOAM [®]	47
4.14	Eigenwerte von \hat{C}_l (links) und der sufficient summary plot der Auftriebskoeffizienten bezüglich des eindimensionalen aktiven Unterrums \hat{w}_l (rechts)	48
4.15	Eigenwerte von \hat{C}_d (links) und der sufficient summary plot der Reibungskoeffizienten bezüglich des eindimensionalen aktiven Unterrums \hat{w}_d (rechts)	48

5.1	Beispielhafte Darstellung des Problems, dass die Orthogonalprojektion eines Punktes $x \in \mathcal{X}$ selbst nicht in \mathcal{X} liegt	53
5.2	Beispielhafte Samples entlang der inaktiven Richtungen zur Konstruktion einer Approximation g der Funktion f	55
5.3	Sufficient summary plot von f_l (blaue Punkte), die Approximation r (rote Gerade) von f_l , der Bereich der Samples, die der tatsächlichen Nebenbedingung an den Auftriebskoeffizienten genügen (grau gefärbt) sowie der Bereich der Samples, die zusätzlich die approximierten Nebenbedingung erfüllen (rot gefärbt)	60
5.4	Zweidimensionaler sufficient summary plot von f_l bezüglich der Richtungen $\hat{w}_{l,1}$ und $\hat{w}_{l,2}$ (die unterschiedlichen Farben der Punkte geben den entsprechenden Funktionswert an) sowie der Bereich der Samples, die der Nebenbedingung an den Auftriebskoeffizienten genügen (grau gefärbt)	61
5.5	Gitter des optimalen Flügelprofils (links) und die Druckverteilung um diesen Flügelquerschnitt (rechts), Abbildung erstellt mit OPENFOAM®	62

Literaturverzeichnis

- [Bel61] BELLMANN, Richard E.: *Adaptive Control Processes: A Guided Tour*. Princeton University Press, 1961 (Princeton Legacy Library)
- [BPDSC17] BAHAMONDE, Sebastian ; PINI, Matteo ; DE SERVI, Carlo ; COLONNA, Piero: Active subspaces for the optimal meanline design of unconventional turbomachinery. In: *Applied Thermal Engineering* 127 (2017), S. 1108 – 1118
- [Caf98] CAFLISCH, Russel E.: Monte Carlo and quasi-Monte Carlo methods. In: *Acta numerica, 1998* Bd. 7. Cambridge Univ. Press, Cambridge, 1998
- [CC02] CHIAROMONTE, Francesca ; COOK, R. D.: Sufficient Dimension Reduction and Graphics in Regression. In: *Annals of the Institute of Statistical Mathematics* 54 (2002), Dec, Nr. 4, S. 768–795
- [CD17] CONSTANTINE, Paul G. ; DOOSTAN, Alireza: Time-dependent global sensitivity analysis with active subspaces for a lithium ion battery model. In: *Statistical Analysis and Data Mining* 10 (2017), Nr. 5, S. 243–262
- [CDW14] CONSTANTINE, Paul G. ; DOW, Eric ; WANG, Qiqi: Active subspace methods in theory and practice: applications to kriging surfaces. In: *SIAM Journal on Scientific Computing* 36 (2014), Nr. 4, S. A1500–A1524
- [CG15] CONSTANTINE, Paul G. ; GLEICH, David: *Computing active subspaces with Monte Carlo*. 2015. – Preprint, arXiv:1408.0545
- [Con15] CONSTANTINE, Paul G.: *Active subspaces. Emerging ideas for dimension reduction in parameter studies*. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM), 2015
- [Coo98] COOK, R. D.: *Regression graphics. Ideas for studying regressions through graphics*. New York, NY: Wiley, 1998
- [Efr79] EFRON, B.: Bootstrap methods: another look at the jackknife. In: *The Annals of Statistics* 7 (1979), Nr. 1
- [Far14] FARIN, Gerald: *Curves and Surfaces for Computer-Aided Geometric Design: A Practical Guide*. Elsevier Science, 2014 (Computer science and scientific computing)
- [Fis22] FISHER, R. A.: On the mathematical foundations of theoretical statistics. In: *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 222 (1922), Nr. 594-604, S. 309–368
- [GC17] GREY, Zachary J. ; CONSTANTINE, Paul G.: *Active subspaces of airfoil shape parameterizations*. 2017. – Preprint, arXiv:1702.02909

- [GT11] GITTENS, Alex ; TROPP, Joel A.: *Tail bounds for all eigenvalues of a sum of random matrices*. 2011. – Preprint, arXiv:1104.4513
- [GTJ⁺21] GALASSI, M. ; THEILER, J. ; JUNGMAN, G. ; GOUGH, B. ; WEITERE: *GNU Scientific Library (GSL)*. <https://www.gnu.org/software/gsl/>. Version: 2.1, Abruf: 28.06.2018
- [GVL96] GOLUB, Gene H. ; VAN LOAN, Charles F.: *Matrix Computations*. The Johns Hopkins University Press, Baltimore und London, 1996
- [HCV07] HADJIANDREOU, M. ; CONEJEROS, Raul ; VASSILIADIS, Vassilis S.: Towards a long-term model construction for the dynamic simulation of HIV infection. In: *Mathematical Biosciences and Engineering* 4 (2007), S. 489–504
- [HJF⁺12] HEATHER, A. ; JANSSENS, M. ; FERRARIS, S. ; OLESEN, M. ; WEITERE: *OpenFOAM®*, *The open source CFD toolbox*. <https://www.openfoam.com/>. Version: v1712, Abruf: 28.06.2018
- [Hul07] HULL, David G.: *Fundamentals of Airplane Flight Mechanics*. Springer-Verlag Berlin Heidelberg, 2007
- [JGCM16] JEFFERSON, Jennifer L. ; GILBERT, James M. ; CONSTANTINE, Paul G. ; MAXWELL, Reed M.: Reprint of: Active subspaces for sensitivity analysis and dimension reduction of an integrated hydrologic model. In: *Computers & Geosciences* 90 (2016), S. 78 – 89. – Uncertainty and Sensitivity in Surface Dynamics Modeling
- [JWP33] JACOBS, Eastman N. ; WARD, Kenneth E. ; PINKERTON, Robert M.: The characteristics of 78 related airfoil sections from tests in the variable-density wind tunnel. 1933 (NACA-TR-460, PB-177874). – NACA Technical Report 460
- [JWZ⁺18] JI, Weiqi ; WANG, Jiaxing ; ZAHM, Olivier ; MARZOUK, Youssef M. ; YANG, Bin ; REN, Zhuyin ; LAW, Chung K.: Shared Low-Dimensional Subspaces for Propagating Kinetic Uncertainty to Multiple Outputs. 190 (2018), 04, S. 146–157
- [Kat95] KATO, Tosio: *Perturbation Theory for Linear Operators*. Springer-Verlag Berlin Heidelberg, 1995
- [LP17] LOUDON, Tyson ; PANKAVICH, Stephen: Mathematical analysis and dynamic active subspaces for a long term model of HIV. In: *Mathematical Biosciences & Engineering* 14 (2017), Nr. 3, S. 709–733
- [LPAC14] LUKACZYK, Trent W. ; PALACOIS, Francisco ; ALONSO, Juan ; CONSTANTINE, Paul G.: Active Subspaces for Shape Optimization. (2014), 01

- [Luk15] LUKACZYK, Trent W.: *Surrogate Modeling and Active Subspaces for Efficient Optimization of Supersonic Aircraft*, Stanford University, dissertation, 2015. <http://adl.stanford.edu/papers/Lukaczyk-Dissertation.pdf>, Abruf: 28.06.2018
- [Mar06] MARSH, Duncan: *Applied Geometry for Computer Graphics and CAD*. Springer London, 2006 (Springer Undergraduate Mathematics Series)
- [Mor13] MORAN, J.: *An Introduction to Theoretical and Computational Aerodynamics*. Dover Publications, 2013 (Dover Books on Aeronautical Engineering)
- [Mur12] MURPHY, Kenneth: *Janeway's Immunobiology*. 8. Garland Science, Taylor & Francis Group, LLC, 2012
- [Nie92] NIEDERREITER, H.: *Random Number Generation and Quasi-Monte Carlo Methods*. Society for Industrial and Applied Mathematics (SIAM), 1992 (CBMS-NSF Regional Conference Series in Applied Mathematics)
- [NW99] NOCEDAL, Jorge ; WRIGHT, Stephen J.: *Numerical Optimization*. Springer New York, 1999 (Springer Series in Operations Research)
- [REw22] RAWLINGS, J. B. ; EKERDT, J. G. ; WEITERE: *GNU Octave*. <https://www.gnu.org/software/octave/>. Version: 4.2.2, Abruf: 28.06.2018
- [Rus10] RUSSI, Trent M.: *Uncertainty Quantification with Experimental Data and Complex System Models*, University of California, Berkeley, dissertation, 2010. <https://cloudfront.escholarship.org/dist/prd/content/qt5d01s25c/qt5d01s25c.pdf>, Abruf: 28.06.2018
- [Sam93] SAMAROV, Alexander M.: Exploring Regression Structure Using Nonparametric Functional Estimation. In: *Journal of the American Statistical Association* 88 (1993), Nr. 423, S. 836–847
- [SC13] SHALIT, Uri ; CHECHIK, Gal: *Efficient coordinate-descent for orthogonal matrices through Givens rotations*. 2013. – Preprint, arXiv:1312.0624
- [Ste73] STEWART, G. W.: Error and Perturbation Bounds for Subspaces Associated with Certain Eigenvalue Problems. In: *SIAM Review* 15 (1973), Nr. 4, S. 727–764
- [SW95] SHWARTZ, Adam ; WEISS, Alan: *Large Deviations For Performance Analysis: Queues, Communication and Computing*. Taylor & Francis, 1995 (Stochastic Modeling Series)
- [TDG⁺18] TEZZELE, Marco ; DEMO, Nicola ; GADALLA, Mahmoud ; MOLA, Andrea ; ROZZA, Gianluigi: *Model Order Reduction by means of Active Subspaces and Dynamic Mode Decomposition for Parametric Hull Shape Design Hydrodynamics*. 2018. – Preprint, arXiv:1803.07377

- [YE03] YAMADA, Isao ; EZAKI, Takato: An Orthogonal Matrix Optimization by dual Cayley Parametrization Technique. In: *4th International Symposium on Independent Component Analysis and Blind Signal Separation (ICA2003)*, April 2003, Nara, Japan (2003)