

**Divide-et-Impera-Verfahren
für das verallgemeinerte
Eigenwertproblem**

Dissertation

zur

**Erlangung des Grades eines Dr.rer.nat.
des Fachbereichs Mathematik
der Fernuniversität-Gesamthochschule-Hagen**

**vorgelegt von
Daniela Keller
aus Herdecke**

Herdecke 1994



Berichterstatter: Prof. Dr. Kresimir Veselić
 Prof. Dr. Hansjörg Linden

Tag der mündlichen Prüfung: 29. August 1994

Inhaltsverzeichnis

1. Einleitung	1
2. Zur Theorie des verallgemeinerten Eigenwertproblems	8
2.1. Grundlegende Begriffe und Aussagen	8
2.2. Vollständige Lösung von Eigenwertproblemen und simultane Diagonalisierung von Matrizenpaaren	16
2.3. Das positiv definite verallgemeinerte Eigenwertproblem	25
3. Einige numerische Verfahren zum verallgemeinerten Eigenwertproblems	29
3.1. Potenzmethode, inverse Iteration und Rayleighquotientenverfahren	29
3.2. Rayleigh-artige Verfahren	30
3.2.1. Die Rayleigh-Ritz-Approximation und die Unterraum-Iteration	
3.2.2. Koordinatenrelaxation	
3.3. Der QZ-Algorithmus	32
3.4. Das verallgemeinerte Jacobi-Verfahren	33
3.5. Weitere Verfahren	34
3.5.1. Bisektionsverfahren	
3.5.2. Lanczos-Algorithmus	
3.6. Verfahren zur Bestimmung des Definitheitsintervalls	35
3.6.1. Ein stochastisches Verfahren	
3.6.2. Ein Koordinatenrelaxationsverfahren zum Test auf positive Definitheit	
4. Der Divide-et-Impera-Algorithmus für ein Matrizenpaar (A, J) mit Tridiagonalmatrix A	41
4.1. Das eigentliche Verfahren	41
4.1.1. Der Divide-Schritt	
4.1.2. Der Impera-Schritt	
4.1.3. Die Lösung des verallgemeinerten Eigenwertproblems $(D - \rho zz^T)y = \lambda Jy$	
4.1.3.1. Minuseigenwerte, Pluseigenwerte und das Definitheitsintervall	
4.1.3.2. Eigenschaften der Funktion w	
4.1.3.3. Deflation	
4.1.3.4. Bestimmung der Eigenvektoren von $(D - \rho zz^T)y = \lambda Jy$	
4.1.4. Die Zerlegung des Eigenwertproblems $Ax = \lambda Jx$	
4.1.5. Testen auf positive Definitheit	

4.2. Die Nullstellensuche	59
4.2.1. Die Bestimmung der Nullstellen der Funktion w	
4.2.2. Das quadratische Einschließungsverfahren	
4.2.3. Praktisches Vorgehen	
4.2.3.1. Ein lineares Einschließungsverfahren	
4.2.3.2. Verschiebungsstrategien	
4.3. J-Orthogonalität von Eigenvektoren und Aussagen zur Genauigkeit des Divide-et-Impera-Verfahrens	72
4.4. Ausblick: Divide-et-Impera-Verfahren für nichtsymmetrische gewöhnliche Eigenwertprobleme	78
5. Zur Reduktion der Matrix A auf Tridiagonalgestalt	81
5.1. Teilweise Reduktion auf Tridiagonalgestalt	81
5.1.1. Die modifizierte Householder-Transformationsmethode	
5.2. Statistische Untersuchungen zur Reduktion von Matrizen auf Tridiagonalform mittels verallgemeinerter Householder-Transformationen	93
6. Divide-et-Impera-Techniken für das verallgemeinerte Eigenwertproblem $Ax = \lambda Jx$	103
6.1. Der Fall einer berandeten Diagonalmatrix	104
6.1.1. Der Deflationsschritt	
6.1.2. Berechnung der Eigenvektoren	
6.1.3. Die Eigenwerte als Nullstellen der Funktion w	
6.1.4. Das Einschließungsverfahren	
6.1.4.1. Einschließung der Nullstellen von w, die im Intervall $(-\infty, \alpha_1)$ liegen	
6.1.4.2. Berechnung der Schranken M_u und M_o	
6.1.4.3. Ein lineares Eingrenzungsverfahren	
6.2. Ein zusammengesetztes Divide-et-Impera-Verfahren für ineinandergeschachtelte berandete Tridiagonalmatrizen	119
Anhang A.	126
A.1. Numerische Ergebnisse	126
A.2. Überblick über die Computerprogramme	145
A.3. Literaturverzeichnis	148

Zusammenfassung

Bekanntlich hat das verallgemeinerte Eigenwertproblem $Ax = \lambda Bx$, wobei A und B reelle symmetrische $n \times n$ -Matrizen sind, zumindest dann n Eigenvektoren, wenn es positiv definit ist, d.h. für geeignete reelle Koeffizienten α, β die Linearkombination $\alpha A + \beta B$ positiv definit wird. Äquivalent dazu ist die Existenz einer (nichtsingulären) Eigenvektormatrix Q, die sowohl A als auch B auf Diagonalgestalt D_A bzw. D_B transformiert:

$$Q^T A Q = D_A, \quad Q^T B Q = D_B$$

In dieser Arbeit wird angenommen, daß B die Gestalt einer Diagonalmatrix J hat, deren Diagonalelemente +1 oder -1 sind. (Dies ist für nichtsinguläres B durch eine Transformation immer erreichbar, die symmetrische Matrix A geht dabei gleichzeitig in eine symmetrische Matrix A_{mod} über.) Für eine Eigenvektormatrix Q gilt dann statt der üblichen Orthogonalität die Beziehung $Q^T J Q = J$.

Auf der Basis des für das Standard eigenwertproblem bekannten Divide-et-Impera-Algorithmus von Bunch, Nielsen, Sorensen u.a. für eine Tridiagonalmatrix und dem Verfahren von Wilkinson für eine berandete Diagonalmatrix werden zwei neue Algorithmen für die entsprechenden verallgemeinerten Eigenwertprobleme entwickelt. Die Nullstellensuche, die einen wesentlichen Teil in beiden Algorithmen darstellt, besteht aus einem speziell hierfür konstruierten neuen Eingrenzungsverfahren und dem bekannten Pegasus-Verfahren.

Mit einer Modifikation des üblichen Householderverfahrens zur Reduktion einer Matrix auf Tridiagonalgestalt wird das Paar (A_{mod}, J) auf die Form (T, J) transformiert: dabei ist T eine Tridiagonalmatrix, wenn während des Transformationsalgorithmus keine Division durch null oder durch eine betragsmäßig sehr kleine Zahl verlangt wird, sonst eine Ineinanderschachtelung von berandeten Tridiagonalmatrizen.

Durch eine Kombination unserer beiden Algorithmen ist es möglich, Paare (T, J) mit T Ineinanderschachtelung berandeter Tridiagonalmatrizen auf positive Definitheit zu untersuchen und gegebenenfalls alle Eigenwerte und Eigenvektoren zu berechnen.

Weiß man von vornherein noch nicht, ob ein Matrizenpaar (A, J) bzw. (T, J) positiv definit ist, so kann man die hier vorgestellten Algorithmen als Testverfahren verwenden; sie entdecken nicht positiv definite Paare im allgemeinen recht frühzeitig.

1. Einleitung

Unter einem *verallgemeinerten Eigenwertproblem* versteht man die Aufgabenstellung, zu quadratischen Matrizen A, B einen Vektor $x \neq 0$ und eine Zahl λ zu finden, die der Beziehung

$$(1.1) \quad Ax = \lambda Bx$$

genügen. Der Vektor x heißt *Eigenvektor* zum *Eigenwert* λ .

Ist die Matrix B in der Eigenwertgleichung nichtsingulär, so läßt sich das verallgemeinerte Eigenwertproblem durch Invertieren von B theoretisch immer auf ein gewöhnliches zurückführen:

$$(1.2) \quad B^{-1}Ax = \lambda x$$

Abgesehen von dem Rechenaufwand, der für die Invertierung von B nötig ist, und den bei schlecht konditioniertem B zu erwartenden Rechenungenauigkeiten, sprechen auch andere Gründe gegen diesen Weg: Häufig sind A und B dünn besetzte Matrizen oder haben eine spezielle Struktur, z.B. Tridiagonalform; die invertierte Matrix B^{-1} - und damit auch $B^{-1}A$ - wird im allgemeinen aber vollbesetzt sein. Auch die Symmetrie von A und B wird nicht zu einer Symmetrie von $B^{-1}A$ führen. Ein Verfahren, das ohne Invertierung auskommt, ist daher für das verallgemeinerte Eigenwertproblem sehr wünschenswert.

Auf verallgemeinerte Eigenwertprobleme stößt man z.B. bei linearen, zeitinvarianten Systemen in der Physik und in der Regelungstechnik, wenn man die zugrundeliegenden linearen Differentialgleichungssysteme

$$(1.3) \quad Ly = Ry$$

(L, R lineare Matrix-Differentialoperatoren bzgl. der Zeit t mit konstanten Koeffizienten) mit Hilfe des Exponentialansatzes $y \equiv y(t) = x e^{\lambda t}$ zu lösen versucht ("übergedämpfte lineare Systeme").

Falls L und R nicht bereits Differentialoperatoren erster Ordnung sind, kann dies in der üblichen Art durch Einführung neuer Variablen für die Ableitungen \dot{y} , \ddot{y} , $\ddot{\ddot{y}}$, ... erreicht werden. Wir dürfen somit immer ein Differentialgleichungssystem der Gestalt

$$(1.4) \quad \left[A_1 + B_1 \frac{d}{dt} \right] y(t) = \left[A_2 + B_2 \frac{d}{dt} \right] y(t)$$

voraussetzen. Dieses geht bei dem angeführten Exponentialansatz in die Matrixbeziehung

$$\left[A_1 - A_2 \right] x = \lambda \left[-B_1 + B_2 \right] x,$$

also in $Ax = \lambda Bx$ mit $A := A_1 - A_2$, $B := -B_1 + B_2$ über.

Bei gekoppelten Systemen in der Mechanik kann B häufig als eine Verallgemeinerung des positiv definiten "Massenoperators"

$$(1.5) \quad M = \begin{bmatrix} m_1 & & 0 \\ & \ddots & \\ 0 & & m_n \end{bmatrix}$$

interpretiert werden, wobei m_1, \dots, m_n die (natürlich positiven) Massen der n wechselwirkenden Teilchen bedeuten.

In der Regelungstechnik ist das zeitliche Verhalten linearer, zeitinvarianter Systeme in der Form

$$L y_{\text{Zustand}} = R y_{\text{Eingang}}$$

beschreibbar. Hängt die Eingangsgröße y_{Eingang} durch eine Rückkopplung

$$y_{\text{Eingang}} = S y_{\text{Zustand}}$$

wieder von der Zustandsgröße y_{Zustand} ab, so erhalten wir ein Differentialgleichungssystem

$$(1.6) \quad L y_{\text{Zustand}} = RS y_{\text{Zustand}}$$

und damit ein verallgemeinertes Eigenwertproblem, wenn man mit dem Exponentialansatz in diese Differentialgleichung geht.

Ein weiteres Gebiet, in dem verallgemeinerte Eigenwertprobleme auftreten können, ist mit der Regelungstechnik verwandt, nämlich die Analyse elektrischer Netzwerke. Auch hier entstehen durch ohmsche, kapazitative und induktive Komponenten (näherungsweise) lineare Systeme, deren zeitliches Verhalten durch Differentialgleichungen der angegebenen Art beschrieben werden kann.

Es ist ziemlich einfach zu beweisen, daß das verallgemeinerte Eigenwertproblem $Ax = \lambda Bx$ im Fall eines positiv definiten, symmetrischen B und eines symmetrischen A n linear unabhängige Eigenvektoren hat. Eine etwas schwächere hinreichende Bedingung ist die positive Definitheit des Paares (A, B). (Das Paar (A, B) heißt positiv definit, wenn für geeignete reelle α, β die Linearkombination $\alpha A + \beta B$ eine positiv definite Matrix ist, vgl. Abschnitt 2.3 und Parlett [46], Golub, van Loan [29]).

Auch wenn die Bedingung der positiven Definitheit eines Matrizenpaares in der Praxis häufig erfüllt ist, so muß doch betont werden, daß sie keineswegs für die Lösbarkeit des zugehörigen Eigenwertproblems notwendig ist. Der QZ-Algorithmus z.B. kann auch auf den Fall eines nicht positiv definiten B angewendet werden, selbst die Symmetrie von A und B ist nicht notwendig. Nachteilig beim QZ-Algo-

ritmus ist aber, daß von vornherein im Komplexen gearbeitet werden muß und auch im Fall eines positiv definiten Paares die Eigenvektoren nicht automatisch reell sind.

Für symmetrisches B kann ein verallgemeinertes Eigenwertproblem $Ax = \lambda Bx$ relativ schnell auf ein Problem der Gestalt $A_{\text{mod}}x = \lambda Jx$, J Diagonalmatrix mit Einsen, Nullen und Minuseinsen auf der Diagonalen, transformiert werden. Ist B zusätzlich nichtsingulär, so läßt sich $J = \text{diag}(\pm 1)$ erreichen. (Ein Algorithmus, der ohne Lösung des Eigenwertproblems für B auskommt, wurde von Brebner und Gard [12] angegeben.) Für das Eigenwertproblem $Ax = \lambda Jx$ stellt das verallgemeinerte Jacobi-Verfahren (Veselič [63]) bei kleineren Dimensionen eine gute Möglichkeit dar, das Paar (A_{mod}, J) auf positive Definitheit zu testen und gegebenenfalls die Eigenwerte und Eigenvektoren zu berechnen. Der wesentliche Vorteil des Jacobi-Verfahrens ist seine Genauigkeit, ein Nachteil aber die geringe Geschwindigkeit (zumindest beim nicht-parallelen Jacobi-Algorithmus) und die Zerstörung einer speziellen Gestalt der Matrix A_{mod} (etwa Bänderstruktur oder dünn besetzt) im Verlauf der Rechnung.

In dieser Arbeit wird als erstes das von Bunch, Nielsen und Sorensen [13] für das normale Eigenwertproblem $Tx = \lambda x$ entwickelte Divide-et-Impera-Verfahren auf den Fall $Tx = \lambda Jx$ verallgemeinert. Wesentliche Voraussetzung in [13] wie zunächst auch bei uns ist die *Tridiagonalität* der Matrix C. Unser Verfahren bricht irgendwann ab - und zwar meist ziemlich früh - wenn das Paar (T, J) nicht positiv definit ist, andernfalls liefert es alle Eigenwerte und Eigenvektoren.

Beim Divide-et-Impera-Verfahren wird das Ausgangsproblem in niedrigerdimensionale Teilprobleme zerlegt, die man direkt lösen kann (*Divide-Schritt*). Aus diesen Lösungen wird dann schrittweise die Gesamtlösung zusammengesetzt (*Impera-Schritt*). Bei unserem Verfahren teilen wir das gegebene Eigenwertproblem solange auf, bis wir direkt lösbare Eigenwertprobleme der Dimension 2 oder 1 erhalten.

Während des Impera-Schrittes müssen wir Eigenwertprobleme lösen, bei denen die Matrix A eine Diagonalmatrix mit Rang-Eins-Störung ist. Ein solches Eigenwertproblem kann zu einem Nullstellenproblem für eine spezielle rationale Funktion w umgeschrieben werden: die Nullstellen von w sind gerade die gesuchten Eigenwerte.

Bunch, Nielsen und Sorensen haben in ihrer Arbeit [13] ein Nullstellenverfahren vorgestellt, das auf einer rationalen Approximation der Funktion w beruht. Die dabei verwendeten Funktionen sind wegen der speziellen Gestalt der Nullstellenfunktion im Fall $J = I$ (streng) monoton, zwischen je zwei Polstellen von w liegt genau eine Nullstelle. Im Fall unseres verallgemeinerten Divide-et-Impera-Verfahrens kann diese Monotonieeigenschaft nicht mehr garantiert werden. Auch die Verteilung der Nullstellen ist für $J \neq I$ anders: alle Nullstellen liegen zwischen dem kleinsten und größten Pol von w , und es gibt ein ausgezeichnetes In-

tervall, das sogenannte Definitheitsintervall, in dem im Fall eines positiv definiten Problems zwei Nullstellen enthalten sind. Das obige Nullstellenverfahren ist daher nicht übertragbar. Aus diesem Grund habe ich ein Eingrenzungsverfahren entwickelt, das das Intervall, in dem die jeweilige Nullstelle oder die zwei Nullstellen liegen müssen, immer weiter verkleinert. Nicht selten ist nach den vier durchgeführten Eingrenzungsschritten (oder schon früher) die Nullstelle bereits gefunden, sonst hat man zumindest ein sehr gutes Startintervall für das Pegasus-Verfahren (eine verbesserte Variante der Regula Falsi), mit dem dann die Nullstelle abschließend berechnet wird. (Falls Nullstelle und Polstelle fast zusammenfallen, wird zusätzlich noch das Bisektionsverfahren als ein äußerst sicheres Verfahren herangezogen.) Dieses Eingrenzungsverfahren nutzt die Tatsache aus, daß w eine sehr spezielle rationale Funktion ist, bei der man genau weiß, wieviele Nullstellen in jedem Intervall liegen - leicht modifiziert, eignet sich das Verfahren aber auch zur Bestimmung von Nullstellen einer beliebigen rationalen Funktion.

Der Fall einer Tridiagonalmatrix A ist nicht nur deshalb interessant, weil solche Matrizen z.B. bei der Diskretisierung von Differentialoperatoren erster oder zweiter Ordnung entstehen, sondern auch, weil jede symmetrische Matrix durch sog. Householdertransformationen auf diese Gestalt gebracht werden kann. Diese Transformationen werden nun auf den Fall eines verallgemeinerten Eigenwertproblems so erweitert, daß die Matrix J (bis auf eventuelle Vertauschungen) unverändert bleibt (vgl. Bunse-Gerstner [15], Steinhoff [54]). Anders als beim normalen Eigenwertproblem kann es jedoch passieren, daß bei diesem Algorithmus durch null oder eine sehr kleine Zahl dividiert werden müßte. Ideen von Dax und Kaniel [20] aufgreifend, lassen wir dann solche kritischen Zeilen und Spalten stehen und machen mit der Transformation der restlichen Matrix weiter. Auf diese Weise bekommen wir statt einer Tridiagonalmatrix eine "Ineinanderschachtelung von berandeten Tridiagonalmatrizen".

Um ein Gefühl dafür zu bekommen, wie häufig Zeilen und Spalten bei der modifizierten Householdermethode *nicht* transformiert werden sollten, wurden hierzu statistische Untersuchungen durchgeführt. Im Einklang mit den angestellten theoretischen Überlegungen trat dieser Fall besonders dann häufig auf, wenn die Matrix J ungefähr gleich viele +1 und -1 hat, die auf der Diagonale einigermaßen gleichmäßig verteilt sind.

"Berandete Diagonalmatrizen" hat Wilkinson in seinem Buch "The Algebraic Eigenvalue Problem" [68] behandelt. Die Lösung des Eigenwertproblems $Cx = \lambda x$ für eine symmetrische, berandete Diagonalmatrix C erfolgt formal ähnlich wie beim Divide-et-Impera-Verfahren: Nach vorausgegangener Deflation sind die Eigenwerte die Nullstellen einer speziellen rationalen Funktion. Die Eigenvektoren zu einfachen Eigenwerten können wie beim Divide-et-Impera-Verfahren direkt durch eine Formel angegeben werden.

Dieses Verfahren habe ich auf ein Problem der Form $Cx = \lambda Jx$ mit J gleich $\text{diag}(\pm 1)$ verallgemeinert. Im Vergleich zum Standardverfahren von Wilkinson ist

die Berechnung der Eigenvektoren bei mehrfachen Eigenwerten etwas aufwendiger, da die berechneten Vektoren noch J -orthogonalisiert werden müssen. Die Probleme bei der Nullstellensuche sind vergleichbar mit denen beim verallgemeinerten Divide-et-Impera-Verfahren: auch hier kann ein Eingrenzungsverfahren entwickelt werden, das analoge Eigenschaften wie das bereits vorgestellte hat. Dieses verallgemeinerte Wilkinson-Verfahren läßt sich leicht auf den Fall erweitern, daß C eine Ineinanderschachtelung von berandeten Diagonalmatrizen ist; theoretisch könnte man es auch auf eine vollbesetzte Matrix C anwenden, aber gut geeignet ist es sicherlich nur für Matrizen mit wenigen Rändern.

Eine Kombination des verallgemeinerten Divide-et-Impera-Verfahrens und des verallgemeinerten Wilkinson-Verfahrens erlaubt jetzt die Lösung des Eigenwertproblems $A_{\text{mod}} x = \lambda Jx$, wobei A_{mod} eine Ineinanderschachtelung von berandeten Tridiagonalmatrizen ist. Zunächst werden Rang-Eins-Abspaltungen wie im Divide-Schritt des Divide-et-Impera-Verfahrens durchgeführt:

$$(1.7) \quad A_{\text{mod}} = A_T - \sum_i |b_i| u_i u_i^T$$

Die Matrix A_T besteht aus entlang der Diagonalen ineinandergeschachtelten Blöcken, die Tridiagonalmatrizen oder berandete Tridiagonalmatrizen sind, die Matrix J wird ebenfalls passend dazu in Blöcke aufgeteilt. Die Teileigenwertprobleme von (A_T, J) können für einen Tridiagonalblock direkt mit dem verallgemeinerten Divide-et-Impera-Verfahren gelöst werden; im Fall eines berandeten Tridiagonalblocks bestimmt man zunächst die vollständige Lösung des inneren verallgemeinerten Tridiagonaleigenwertproblems, dann gelangt man mittels geeigneter Transformationen auf ein verallgemeinertes Eigenwertproblem mit berandeter Diagonalmatrix. Dieses Problem löst man mit unserem modifizierten Wilkinson-Verfahren. Die Lösungen der Teileigenwertprobleme werden anschließend durch die Impera-Schritte miteinander verbunden, wodurch sich am Ende die Lösung des Gesamteigenwertproblems ergibt. Eine solche Lösung existiert sicherlich dann, wenn das Paar (A_{mod}, J) positiv definit ist; unser Verfahren stoppt, sobald ein Teilproblem diese Bedingung verletzt.

Das verallgemeinerte Divide-et-Impera-Verfahren liefert meist sehr gute Ergebnisse. Bei einigen sehr speziellen Matrizen (wie z.B. den sogenannten Wilkinson-Matrizen) kann es allerdings zu einem teilweisen Verlust der (J) -Orthogonalität der Eigenvektoren kommen. Dies ist ein schon für das Standard-Divide-Verfahren bekanntes Phänomen. (Es tritt auf, wenn zwei Nullstellen der rationalen Funktion w fast vollständig mit dem zwischen ihnen liegenden Pol übereinstimmen.) Mit dieser Schwierigkeit könnte man wahrscheinlich fertigwerden, indem man die folgenden, beim Standard-Divide-Verfahren bereits erprobten Modifikationen benutzt:

1. Kritische Terme in der Nullstellensuche und der Eigenvektorformel werden mit höherer Genauigkeit ausgewertet (vgl. Sorensen, Tang [53]).
2. Berechnung der Eigenvektoren mit inverser Iteration (vgl. Gates [27]).
3. Bestimmung der Eigenvektoren mittels einer alternativen Formel (vgl. Gu und Eisenstat [32]).

In dieser Arbeit wird nur kurz auf diese Verfahren eingegangen, in der aktuellen Implementierung des Algorithmus wird die Möglichkeit 1 durch Verwendung von Single/Double-Variablen getestet. Ein Vergleich der drei Lösungsansätze für unser verallgemeinertes Divide-et-Impera-Verfahren wäre sicher interessant.

Theoretische Fehlerabschätzungen und Genauigkeitsaussagen sind für das verallgemeinerte Divide-et-Impera-Verfahren noch nicht durchgeführt worden. Die bekannten Ergebnisse aus den Arbeiten von Bunch, Nielsen, Sorensen [13], Cuppen [19] und Sorensen, Tang [53] lassen sich nicht direkt übertragen: Die Eigenvektoren sind im Spezialfall $J = I$ orthogonal, während wir mit J -orthogonalen Vektoren arbeiten, deren euklidische Norm keineswegs 1 sein muß; desweiteren kann für $J = I$ die Ableitung der Nullstellenfunktion an den Nullstellen nach unten abgeschätzt werden, was für $J \neq I$ nicht mehr möglich ist, da die Ableitung der Nullstellenfunktion dort theoretisch beliebige Werte aus \mathbb{R} annehmen kann. Gerade diese Abschätzung der Ableitung ist aber für die Beweise besonders wichtig.

Da wir in unserem Divide-et-Impera-Verfahren das Eigenwertproblem so lange aufteilen, bis wir Teilprobleme der Dimension 1 oder 2 erhalten, ist das Nichterfülltsein der positiven Definitheit oft schon bei niedrigdimensionalen Teilproblemen erkennbar. Eine etwas andere - für eine rekursive Programmierung günstigere - Möglichkeit bestünde darin, nur so lange aufzuteilen, bis man Eigenwertprobleme hat, auf die sich z.B. das verallgemeinerte Jacobi-Verfahren (vgl. Veselič [63]) oder das QZ-Verfahren (vgl. Golub/van Loan [29]) anwenden lassen.

Unklar ist, ob und wie das verallgemeinerte Divide-et-Impera-Verfahren auf den Fall eines nicht positiv definiten Paares (T, J) verallgemeinerbar ist. -

Diese Arbeit baut sich folgendermaßen auf:

Im Kapitel 2 beschäftigen wir uns mit der theoretischen Lösbarkeit des verallgemeinerten Eigenwertproblems $A x = \lambda B x$ und leiten hierzu einige prinzipielle Ergebnisse ab. Die meisten dieser Aussagen dürften in etwa bekannt sein, sind aber so noch nicht zusammengefaßt worden.

Im Kapitel 3 stellen wir zunächst bekannte Verfahren zur Lösung des symmetrischen verallgemeinerten Eigenwertproblems $A x = \lambda B x$ vor; die meisten dieser Methoden verlangen die positive Definitheit von B . Das verallgemeinerte Jacobi-Verfahren dagegen kommt mit der positiven Definitheit des Paares (A, J) aus. Sehr universell ist das QZ-Verfahren, wenn es im Komplexen formuliert wird; es versucht, für beliebige Matrizenpaare (A, B) die Eigenwertbestimmung durchzuführen.

Unabhängig von dem Divide-et-Impera-Verfahren, das den Hauptteil dieser Arbeit bildet, wurden zwei einfachere Algorithmen entwickelt, die jedoch das Paar (A, B) nur auf positive Definitheit testen und fallweise ein μ aus \mathbb{R} liefern, für das $A - \mu B$ positiv definit wird. Diese Algorithmen werden kurz in Abschnitt 3.6 skizziert. Der erste stellt eine Art Monte-Carlo-Optimierung dar, der zweite beruht auf einem Theorem von Kovač-Veselič [38] über Spurminimierung. Kennt man ein solches μ , so kann man zu einem Paar (A, B') mit einem positiv definiten B' der Gestalt $B' := A - \mu B$ übergehen und das Eigenwertproblem $Ax = \lambda B'x$ mit einem der dafür geeigneten Verfahren lösen.

Im Kapitel 4 - das den Kern dieser Arbeit bildet - stellen wir ausführlich unser verallgemeinertes Divide-et-Impera-Verfahren für das Paar (T, J) - mit T Tridiagonalmatrix und $J = \text{diag}(\pm 1)$ - vor.

Im Kapitel 5 beschreiben wir die (evtl. nur teilweise) Reduktion einer Matrix auf Tridiagonalgestalt mittels modifizierter Householder-Transformationen. Um ein Gefühl zu bekommen, wie häufig der Fall einer unvollständigen Reduktion auftritt, führen wir neben theoretischen statistischen Untersuchungen auch Simulationen unter Verwendung von Pseudozufallszahlen durch.

In Kapitel 6 wird zunächst ein Algorithmus zur Lösung des speziellen Eigenwertproblems $C x = \lambda J x$ - mit C berandete Diagonalmatrix - beschrieben. Eine Kombination dieses Algorithmus mit dem verallgemeinerten Divide-et-Impera-Verfahren aus Kapitel 4 erlaubt uns dann, Eigenwertprobleme $T x = \lambda J x$ auch dann zu lösen, wenn T eine Ineinanderschachtelung berandeter Tridiagonalmatrizen ist.

Ein Anhang bringt eine systematische Zusammenstellung und Diskussion aller gewonnenen Computer-Ergebnisse. Diese wurden auf PCs mit den INTEL-Prozessoren 80386/387 und 80486, auf SUN-RISC-Workstations und auf dem Siemens-Vektorrechner SNI S-600/20 erhalten. Die Programme sind in ihrer ursprünglichen Form in Turbo-Pascal (für große Matrizen ist der "Protected Mode" in Version 7.0 notwendig) geschrieben und wurden später in C übertragen. Die wesentlichen Teile des Turbo-Pascal-Programms werden in einem weiteren Anhang präsentiert.

2. Zur Theorie des verallgemeinerten Eigenwertproblems

2.1 Grundlegende Begriffe und Aussagen

Es wird geprüft, wie weit sich die bekannten Begriffe und Aussagen zum normalen Eigenwertproblem auf verallgemeinerte Eigenwertprobleme erweitern lassen. Das neue - beim normalen Eigenwertproblem nicht auftretende - Phänomen "singulärer Büschel" wird ausführlicher untersucht.

Im Hinblick auf die später nur im Reellen funktionierenden Algorithmen werden wir für das verallgemeinerte Eigenwertproblem

$$(2.1) \quad A x = \lambda B x$$

in dieser Arbeit stets den Vektorraum \mathbb{R}^n (über dem Körper \mathbb{R} der reellen Zahlen) annehmen; der Eigenwert soll aus dem zugrundegelegten Körper, also bei uns reell sein. Die Übertragung auf den \mathbb{C}^n über dem Körper \mathbb{C} der komplexen Zahlen macht aber größtenteils keine Schwierigkeiten.

Anders als beim "normalen" Eigenwertproblem

$$(2.1') \quad A x = \lambda x,$$

das dem Fall $B = I$, I n -dimensionale Einheitsmatrix, entspricht, ist sogar eine Verallgemeinerung auf den Fall nichtquadratischer Matrizen möglich, denn die Gleichung (2.1) macht auch noch Sinn, wenn A und B (m, n)-Matrizen sind. Dies soll jedoch in dieser Arbeit nicht weiterverfolgt werden.

Natürlich kann im Fall einer nichtsingulären Matrix B das verallgemeinerte Eigenwertproblem theoretisch immer durch Multiplikation beider Seiten mit B^{-1} auf ein gewöhnliches zurückgeführt werden. Es wurde aber bereits betont, daß dies in der Regel nicht empfehlenswert ist: vor allem in der Mechanik sind die Matrizen A und B symmetrisch, die Matrix $B^{-1} A$ wird jedoch normalerweise keinesfalls symmetrisch sein. Wie das folgende Lemma zeigt, sind (nichtsymmetrische) gewöhnliche Eigenwertprobleme sogar äquivalent zu verallgemeinerten Eigenwertproblemen mit nichtsingulärem B :

Lemma 2.1: Jede (reelle) Matrix M kann als Produkt $M = A B^{-1}$ mit A, B symmetrisch und B invertierbar geschrieben werden.

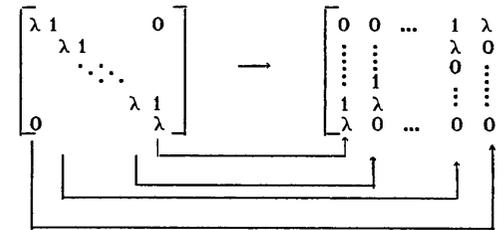
Beweis: (Eine Beweisskizze mit etwas anderen Überlegungen findet sich im Buch von Parlett [46])

Mit Hilfe einer reellen, nichtsingulären Transformationsmatrix F kann M auf Jordan-Normalform M_{Jordan} gebracht werden:

$$M_{\text{Jordan}} = F^{-1} M F \quad \text{mit}$$

$$(2.2) \quad M_{\text{Jordan}} = \begin{bmatrix} J_1 & & 0 \\ & J_2 & \\ & & \ddots \\ 0 & & & J_m \end{bmatrix} \quad \text{und} \quad J_r = \begin{bmatrix} \lambda_r & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_r \end{bmatrix}.$$

Ein "Jordankästchen" J_r der Dimension k läßt sich aber durch Vertauschen der Spalten in eine symmetrische Matrix verwandeln:



Diese Umkehrung der Spaltenreihenfolge kann aber auch durch eine Matrix-Operation erzeugt werden, denn es gilt für beliebige Vektoren $a_1, \dots, a_k \in \mathbb{R}^k$:

$$[a_1, a_2, \dots, a_k] \begin{bmatrix} 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & & 1 & 0 \\ \vdots & & \ddots & \vdots & \\ 0 & 1 & & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \end{bmatrix} = [a_k, a_{k-1}, \dots, a_1].$$

Zusammensetzen der Kästchen zur ursprünglichen Matrix führt zu

$$M_{\text{Jordan}} = \begin{bmatrix} J_1 & & \\ & J_2 & \\ & & \ddots \\ & & & J_m \end{bmatrix} = \begin{bmatrix} A_1 & & \\ & A_2 & \\ & & \ddots \\ & & & A_m \end{bmatrix} \begin{bmatrix} P_1 & & \\ & P_2 & \\ & & \ddots \\ & & & P_m \end{bmatrix}$$

$$\text{mit } A_r = \begin{bmatrix} 0 & & & 1 & \lambda_r \\ & \ddots & & \vdots & \\ & & \ddots & \vdots & \\ \lambda_r & & & & 0 \end{bmatrix} \quad \text{und} \quad P_r = \begin{bmatrix} 0 & & & 1 \\ & \ddots & & \\ & & \ddots & \\ 1 & & & 0 \end{bmatrix}.$$

Die Matrix $P = \begin{bmatrix} P_1 & & 0 \\ & \ddots & \\ 0 & & P_r \end{bmatrix}$ ist symmetrisch und hat als Determinante den

Wert +1 oder -1, also existiert die Inverse.

Die Matrix $A = \begin{bmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_r \end{bmatrix}$ ist ebenfalls symmetrisch.

Somit gilt insgesamt

$$M = F M_{\text{Jordan}} F^{-1} = F A P F^{-1} = F A F^T (F^T)^{-1} P F^{-1} \\ = A' B_1' \text{ mit } A' = F A F^T \text{ und } B_1' = (F^{-1})^T P F^{-1}.$$

Die Matrizen A', B_1' sind offensichtlich symmetrisch. Das gleiche gilt auch für die Inverse $B' := B_1'^{-1}$. Damit haben wir die angegebene Darstellung $M = A' B'^{-1}$ gewonnen.

Q.E.D.

Gehen wir bei der hier definierten Faktorisierung nicht von M , sondern von M^T aus, so erhalten wir eine entsprechende Darstellung in umgekehrter Reihenfolge:

$$M = (M^T)^T = (A' B'^{-1})^T = (B'^{-1})^T A'^T = (B'^T)^{-1} A'^T = B^{-1} A.$$

Ein (beliebiges nichtsymmetrisches) Eigenwertproblem $A x = \lambda x$ kann daher stets als ein *symmetrisches verallgemeinertes Eigenwertproblem* aufgefaßt werden:

$$(2.3) \quad A x = \lambda x \iff A_2^{-1} A_1 x = \lambda x \iff A_1 x = \lambda A_2 x$$

mit A_1, A_2 symmetrisch.

Aus dem Lemma (2.1) folgt auch sofort, daß wir bei einem verallgemeinerten Eigenwertproblem die Matrix auf der rechten Seite (oder auf der linken Seite) immer als symmetrisch annehmen dürfen:

$$(2.4) \quad A x = \lambda B x \iff A x = \lambda B_1^{-1} B_2 x \iff B_1 A x = \lambda B_2 x,$$

wobei hier B_2 symmetrisch ist.

Ein über das nichtsymmetrische, gewöhnliche Eigenwertproblem hinausgehender Gesichtspunkt tritt in Erscheinung, falls die Matrix B auf der rechten Seite singular ist: Dann ist auch die symmetrische Matrix B_2 singular, und eine Diagonalisierung von B_2 durch eine orthogonale Matrix U führt (evtl. nach geeigneter Permutation) auf die Darstellung

$$B_2 := U \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} U^T \text{ mit } D = \text{diag}(\lambda_1, \dots, \lambda_m), \lambda_1, \dots, \lambda_m \neq 0.$$

Durch Multiplikation dieser Diagonalmatrix mit

$$S_1 = \begin{bmatrix} \frac{1}{\sqrt{|\lambda_1|}} & & 0 \\ & \ddots & \\ & & \frac{1}{\sqrt{|\lambda_m|}} \\ 0 & & & I_{n-m} \end{bmatrix}, \quad S_2 = \begin{bmatrix} \frac{\text{sgn}(\lambda_1)}{\sqrt{|\lambda_1|}} & & 0 \\ & \ddots & \\ & & \frac{\text{sgn}(\lambda_m)}{\sqrt{|\lambda_m|}} \\ 0 & & & I_{n-m} \end{bmatrix},$$

von links bzw. von rechts - "sgn" bedeutet die übliche Vorzeichenfunktion - bekommen wir schließlich

$$A x = \lambda B x \iff B_1 A x = \lambda B_2 x \iff U^T B_1 A (U U^T) x = \lambda U^T B_2 (U U^T) x \\ \iff U^T B_1 A U U^T x = \lambda \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} U^T x \\ \iff \underbrace{S_1 U^T B_1 A U S_2}_{=: \tilde{A}} S_2^{-1} U^T x = \lambda \begin{bmatrix} I_m & 0 \\ 0 & 0 \end{bmatrix} S_2^{-1} U^T x.$$

Da $S_2^{-1} U^T x$ für $x \in \mathbb{R}^n$ den ganzen \mathbb{R}^n durchläuft, ist ein verallgemeinertes Eigenwertproblem immer einem der Form

$$(2.5) \quad \tilde{A} x = \lambda \begin{bmatrix} I_m & 0 \\ 0 & 0 \end{bmatrix} x$$

äquivalent, wobei das m gleich dem Rang von B ist.

Wir werden im nächsten Abschnitt sehen, daß für $m < n$ im allgemeinen nicht mehr n Eigenvektoren gefunden werden können.

Vom Standpunkt der Numerik ist bemerkenswert, daß der Fall eines singulären B häufig beseitigt werden kann. Für ein beliebiges σ gilt nämlich offensichtlich:

$$(2.6) \quad A x = \lambda B x \iff (A + \lambda \sigma A) x = \lambda (B + \sigma A) x \\ \iff A x = \lambda \frac{1}{1 + \lambda \sigma} (B + \sigma A) x$$

Sofern nicht $B + \sigma A$ für jedes σ singular ist, kann für ein geeignetes σ_0 die Matrix $(B + \sigma_0 A)$ verhältnismäßig risikolos hinsichtlich der Genauigkeit invertiert und das neue (äquivalente) Eigenwertproblem $A x = \lambda B' x$ betrachtet

werden. (Ist A invertierbar, so mag auch das Eigenwertproblem $\lambda^{-1} \mathbf{x} = A^{-1} B \mathbf{x}$ für $\lambda \neq 0$ untersucht werden. Der Eigenwert $\lambda = 0$ darf bei nichtsingulärem A natürlich nicht vorkommen!)

Faßt man die angestellten Überlegungen zusammen, so erkennt man, daß - anders als beim üblichen Eigenwertproblem - ein verallgemeinertes Eigenwertproblem meist in ein symmetrisches transformiert werden kann:

Falls $A + \sigma_0 B$ für irgendein σ_0 invertierbar ist, betrachtet man das nichtsinguläre Problem

$$A \mathbf{x} = \lambda B' \mathbf{x} \quad \text{mit} \quad B' = A + \sigma_0 B.$$

Dieses ist aber äquivalent zu $B^{-1} A \mathbf{x} = \lambda \mathbf{x}$.

Die Matrix $M := B^{-1} A$ kann nun nach Lemma 2.1 in der Form $M = B_1^{-1} A_1$ mit symmetrischen Matrizen A_1, B_1 geschrieben werden. Mit diesen geht das Eigenwertproblem in die Form $A_1 \mathbf{x} = \lambda B_1 \mathbf{x}$ eines *symmetrischen* verallgemeinerten Eigenwertproblems über.

Die entscheidende Bedingung, daß die Schar $A + \sigma B$ nicht für *alle* $\sigma \in \mathbb{R}$ singular ist, kann man symmetrischer formulieren:

Ein Matrizenpaar (A, B) heißt *singular*, wenn das "Büschel" $\{\alpha A + \beta B \mid \alpha, \beta \in \mathbb{R}\}$ ausschließlich aus singulären Matrizen besteht, sozusagen ein *singuläres Büschel* ist, sonst heißt es nichtsingulär.

(Ist $\alpha A + \beta B$ invertierbar, so auch $A + \sigma B$ mit $\sigma := \beta/\alpha$, sofern $\alpha \neq 0$. Im Sonderfall $\alpha = 0$ muß B selbst invertierbar sein. Dann kann aber

$$P(\sigma) := \det(A + \sigma B) = \sigma^n \left[\det\left(\frac{1}{\sigma} A + B\right) \right],$$

das ein Polynom n-ten Grades darstellt, nicht identisch null sein, sonst würde dies auch für $P(\sigma)/\sigma^n = \det(\frac{1}{\sigma} A + B)$ und damit für den Limes $\sigma \rightarrow \infty$, also für $\det(B)$ gelten.)

Falls $\text{Ker } A \cap \text{Ker } B \neq 0$, so ist $B + \sigma A$ stets singular. Da aber dieser gemeinsame Unterraum - zumindest bei symmetrischen Matrizen A, B - abspaltbar ist (s. später), erweist sich dieser Fall als nicht so kritisch. Leider kann es aber auch passieren, daß A und B nur den Nullvektor als gemeinsamen Kern haben, und trotzdem enthält die Matrizenschar

$$\{B + \sigma A \mid \sigma \in \mathbb{R}\} \quad (\text{bzw. allgemeiner das "Büschel" } \{\alpha A + \beta B \mid \alpha, \beta \in \mathbb{R}\})$$

keine nichtsinguläre Matrix, wie die folgenden Beispiele zeigen.

Im nichtsymmetrischen Fall ist dies bereits für $n = 2$ möglich:

Sei
$$A := \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, \quad B := \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Dann ist $\text{Ker } A = \left\{ \lambda \begin{pmatrix} 1 \\ 1 \end{pmatrix} \mid \lambda \in \mathbb{R} \right\}$ und $\text{Ker } B = \left\{ \lambda \begin{pmatrix} 0 \\ 1 \end{pmatrix} \mid \lambda \in \mathbb{R} \right\}$,

also $\text{Ker } A \cap \text{Ker } B = 0$. Andererseits gilt

$$\det(A + \sigma B) = \begin{vmatrix} 1 + \sigma & -1 \\ 0 & 0 \end{vmatrix} = 0,$$

d.h. $A + \sigma B$ ist für alle σ singular.

Im symmetrischen Fall liefern die folgenden (3x3)-Matrizen ein Beispiel für den gleichen Sachverhalt:

$$A := \begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \quad B := \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Wie man schnell nachprüft, gilt nämlich

$$\text{Ker } A = \left\{ \lambda \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \mid \lambda \in \mathbb{R} \right\}, \quad \text{Ker } B = \left\{ \lambda \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \mid \lambda \in \mathbb{R} \right\}$$

und damit $\text{Ker } A \cap \text{Ker } B = 0$.

Wegen

$$\det(A + \sigma B) = \begin{vmatrix} 1 + \sigma & 2 & 1 \\ 2 & 3 - \sigma & 1 \\ 1 & 1 & 0 \end{vmatrix} = 1 \cdot (2 - 3 + \sigma) - 1 \cdot (1 + \sigma - 2) = 0$$

ist aber $A + \sigma B$ immer singular.-

Das Phänomen einer stets singulären Matrizenschar $A + \sigma B$ ist auch aus einem anderen Grund interessant:

Wie im Fall des gewöhnlichen Eigenwertproblems kann das *charakteristische Polynom*

$$(2.7) \quad P(\lambda) := \det(A - \lambda B)$$

(mit B statt der Einheitsmatrix I) zur Lösung des Eigenwertproblems verwendet werden. Eine Zahl λ ist genau dann Eigenwert, wenn sie Nullstelle des charakteristischen Polynoms ist.

Der Grad des charakteristischen Polynoms muß jedoch nicht mehr gleich n sein:

Satz 2.2: Für beliebige quadratische Matrizen A, B gilt: $\text{grad } P \leq \text{Rang}(B)$

Beweis: Wir haben in Lemma 2.1 gesehen, daß wir $B = B_1^{-1} B_2$ schreiben und B_2 durch nichtsinguläre Matrizen F_1, F_2 auf die Blockdiagonalgestalt

$$\text{diag}(I_m, 0) \quad \text{mit} \quad m = \text{Rang}(B_2) = \text{Rang}(B)$$

bringen können.

Also gilt:

$$\begin{aligned} \det(A - \lambda B) &= (\det(B_1))^{-1} \det(B_1 A - \lambda B_2) \\ &= (\det(B_1))^{-1} \det\left(F_1^{-1} \left(\underbrace{F_1 B_1 A F_2^{-1}}_{=: A'} - \lambda \begin{bmatrix} I_m & 0 \\ 0 & 0 \end{bmatrix}\right) F_2\right) \\ &= (\det(B_1))^{-1} (\det(F_1))^{-1} \det(F_2) \det\left(A' - \lambda \begin{bmatrix} I_m & 0 \\ 0 & 0 \end{bmatrix}\right). \end{aligned}$$

Aus

$$\begin{vmatrix} a'_{11} - \lambda & a'_{12} & \dots & a'_{1n} \\ & \ddots & & \\ \vdots & & a'_{mm} - \lambda & \\ & & & a'_{m+1,m+1} \\ & & & \ddots & \\ a'_{n1} & a'_{n2} & \dots & & a'_{nn} \end{vmatrix} = (-\lambda)^m a'_{m+1,m+1} \cdot \dots \cdot a'_{nn} + O(\lambda^{m-1})$$

liest man sofort die Richtigkeit der Aussage ab.

Q.E.D.

Nach den vorhergehenden Gegenbeispielen kann es also passieren, daß das charakteristische Polynom identisch verschwindet und damit jedes λ Eigenwert ist. Es ist aber auch der Fall nicht ausgeschlossen, daß P gleich einer nichtverschwindenden Konstante ist: in diesem Fall existiert überhaupt kein Eigenwert!

Nach diesen doch etwas überraschenden Fakten kommen wir jetzt zu einigen Eigenschaften, die bereits vom üblichen Eigenwertproblem bekannt sind.

Satz 2.3: Es seien $\mathbf{x}_1, \dots, \mathbf{x}_r$ Eigenvektoren zu verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_r$. Falls $\text{Ker } B = \mathbf{0}$ gilt, sind $\mathbf{x}_1, \dots, \mathbf{x}_r$ linear unabhängig.

Beweis:

Es seien (o. E.) $\mathbf{x}_1, \dots, \mathbf{x}_k$ ($k < r$) linear unabhängig, aber $\mathbf{x}_1, \dots, \mathbf{x}_{k+1}$ linear abhängig. Dann gilt

$$\mathbf{x}_{k+1} = \alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k.$$

Wenden wir auf beide Seiten den Operator $(A - \lambda_{k+1} B)$ an, so erhalten wir

$$\begin{aligned} \mathbf{0} &= \alpha_1 (A - (\lambda_1 + \lambda_{k+1} - \lambda_1) B) \mathbf{x}_1 + \dots + \alpha_k (A - (\lambda_k + \lambda_{k+1} - \lambda_k) B) \mathbf{x}_k \\ &= \alpha_1 (\lambda_1 - \lambda_{k+1}) B \mathbf{x}_1 + \dots + \alpha_k (\lambda_k - \lambda_{k+1}) B \mathbf{x}_k \\ &= B [\alpha_1 (\lambda_1 - \lambda_{k+1}) \mathbf{x}_1 + \dots + \alpha_k (\lambda_k - \lambda_{k+1}) \mathbf{x}_k]. \end{aligned}$$

Da nach Voraussetzung der Kern von B nur aus dem Nullvektor besteht, gilt

$$\alpha_1 (\lambda_1 - \lambda_{k+1}) \mathbf{x}_1 + \dots + \alpha_k (\lambda_k - \lambda_{k+1}) \mathbf{x}_k = \mathbf{0}.$$

Die Faktoren $(\lambda_i - \lambda_{k+1})$, $i = 1, \dots, k$ sind ungleich null, mindestens ein α -Koeffizient ist ebenfalls ungleich null, also sind $\mathbf{x}_1, \dots, \mathbf{x}_k$ linear abhängig im Gegensatz zur Annahme.

Q.E.D.

Falls B einen nichttrivialen Kern hat, ist es dagegen möglich, daß ein und derselbe Vektor Eigenvektor zu unendlich vielen Eigenwerten ist:

Wenn nämlich $\text{Ker } A \cap \text{Ker } B \neq \mathbf{0}$ ist, so gilt für jedes nichtverschwindende \mathbf{x} aus $\text{Ker } A \cap \text{Ker } B$

$$A \mathbf{x} = \mathbf{0} \quad \text{und} \quad B \mathbf{x} = \mathbf{0}, \quad \text{d.h.} \quad A \mathbf{x} = \lambda B \mathbf{x} \quad \text{für alle } \lambda \in \mathbb{R} \text{ (oder } \mathbb{C}).$$

Vom üblichen *symmetrischen* Eigenwertproblem weiß man, daß Eigenvektoren zu verschiedenen Eigenwerten nicht nur linear unabhängig, sondern darüber hinaus auch orthogonal sind. Die Übertragung auf verallgemeinerte symmetrische Eigenwertprobleme lautet:

Satz 2.4: Zwei Eigenvektoren $\mathbf{x}_1, \mathbf{x}_2$ zu zwei verschiedenen Eigenwerten des symmetrischen Eigenwertproblems $A \mathbf{x} = \lambda B \mathbf{x}$ sind "B-orthogonal", d.h. es gilt

$$(\mathbf{x}_1, \mathbf{x}_2)_B = 0,$$

wobei $(\mathbf{x}, \mathbf{y})_B$ das (im allgemeinen) nicht positiv definite "Skalarprodukt" $\mathbf{x}^T B \mathbf{y} \equiv (\mathbf{x}, B \mathbf{y})$ bedeutet.

Beweis:

Aus

$$(\mathbf{x}_1, A \mathbf{x}_2) = (\mathbf{x}_1, \lambda_2 B \mathbf{x}_2) = \lambda_2 (\mathbf{x}_1, B \mathbf{x}_2)$$

und

$$(\mathbf{x}_1, A \mathbf{x}_2) = (A^T \mathbf{x}_1, \mathbf{x}_2) = (A \mathbf{x}_1, \mathbf{x}_2) = (\lambda_1 B \mathbf{x}_1, \mathbf{x}_2) = \lambda_1 (\mathbf{x}_1, B \mathbf{x}_2)$$

folgt durch Subtraktion

$$(\lambda_2 - \lambda_1) (\mathbf{x}_1, B \mathbf{x}_2) = 0, \quad \text{also} \quad (\mathbf{x}_1, B \mathbf{x}_2) = 0. \quad \text{Q.E.D.}$$

Da - wie bereits erwähnt - ein nichtsymmetrisches verallgemeinertes Eigenwertproblem meist auf ein symmetrisches zurückgeführt werden kann, ist die Aussage von Satz 2.4 recht nützlich.

2.2 Vollständige Lösung von Eigenwertproblemen und simultane Diagonalisierung von Matrixpaaren

Es wird gezeigt, daß die für das gewöhnliche Eigenwertproblem gültige Äquivalenz zwischen vollständiger Lösung des Eigenwertproblems und Diagonalisierbarkeit der Matrix beim verallgemeinerten Eigenwertproblem nur unter Zusatzvoraussetzungen - wie Nichtsingularität von B oder Symmetrie von A und B - richtig ist. Für die theoretische Formulierung erweist sich der Begriff der Diagonalisierbarkeit als günstig.

Beim gewöhnlichen Eigenwertproblem $Ax = \lambda x$ im \mathbb{R}^n ist derjenige Fall besonders interessant, in dem n linear unabhängige Eigenvektoren x_1, \dots, x_n existieren. Faßt man diese n (Spalten-) Vektoren zu einer Matrix X zusammen und bildet mit den Eigenwerten $\lambda_1, \dots, \lambda_n$ eine Diagonalmatrix $\Lambda := \text{diag}(\lambda_1, \dots, \lambda_n)$, so gilt

$$A X = X \Lambda.$$

Da die Matrix X nach Voraussetzung maximalen Rang hat, ist sie invertierbar, und es gilt

$$X^{-1} A X = \Lambda,$$

d.h. A kann durch die Matrizen $F := X^{-1}$, $G := X$ diagonalisiert werden.

Wir werden jetzt eine Übertragung auf das verallgemeinerte Eigenwertproblem versuchen.

Definition 2.5: Das verallgemeinerte Eigenwertproblem $Ax = \lambda Bx$, A, B nxn-Matrizen, heie vollständig lösbar, wenn es n linear unabhängige Eigenvektoren x_1, \dots, x_n gibt.

Da der Grad des charakteristischen Polynoms $P(\lambda) = \det(A - \lambda B)$ kleiner als n sein kann, ist es für die vollständige Lösbarkeit des verallgemeinerten Eigenwertproblems nicht mehr ausreichend, wenn bei allen Eigenwerten geometrische und algebraische Vielfachheit übereinstimmen, d.h. kein Eigenwert defektiv ist. Deswegen wollen wir hier die Begriffe "defektiver Eigenwert" und "defektive Matrix" vermeiden.

Fassen wir die Spaltenvektoren x_1, \dots, x_n zur "Eigenvektormatrix" X zusammen, so bedeutet vollständige Lösbarkeit das Bestehen der Matrix-Gleichung

$$(2.8) \quad A X = B X \Lambda,$$

wobei $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ die "Eigenwertmatrix" darstellt.

Ist A symmetrisch, so sind die Eigenvektoren x_1, \dots, x_n beim Standardeigenwertproblem orthogonal bzw. können im Entartungsfall orthogonal gewählt werden; nach Normierung der Eigenvektoren wird X zu einer orthogonalen Matrix, weshalb wir auch

$$X^T A X = \Lambda$$

schreiben dürfen: A ist der Diagonalmatrix Λ "unitär" (orthogonal)-äquivalent.

Die Orthogonalitätseigenschaft der Eigenvektormatrix beim (gewöhnlichen) symmetrischen Eigenwertproblem kann in folgender Weise auf das verallgemeinerte Eigenwertproblem übertragen werden:

Wir nehmen dabei zur Vereinfachung an, daß B nichtsingular ist und damit auf eine Diagonalmatrix J transformiert werden kann, deren Diagonalelemente +1 oder -1 sind, so daß $J^2 = I$ erfüllt ist. Für Eigenvektoren x_i, x_j zu verschiedenen Eigenwerten λ_i, λ_j gilt

$$(x_i, J x_j) = x_i^T J x_j = 0.$$

Für Eigenvektoren zu gleichen Eigenwerten können wir - ähnlich wie beim Gram-Schmidt-Verfahren - eine Linearkombination verwenden, um J-Orthogonalität zu erreichen. Da nach dem Sylvesterschen Trägheitssatz die Anzahl der Plus - und Minuseinsen gleich bleibt, gilt

$$(x_i, J x_i) = x_i^T J x_i > 0 \text{ oder } < 0$$

und nach eventueller Permutation und Normierung für die Eigenvektormatrix X:

$$(2.9) \quad X^T J X = J, \text{ woraus durch Multiplikation } J X^T J X = J^2 = I$$

folgt. Die nichtsinguläre Matrix $X^* := J X^T J$ ist also invers zu X oder anders ausgedrückt, X ist "J-orthogonal" :

$$\langle u, X v \rangle = u^T J X v = u^T J X J J v = (J X^T J u)^T J v = \langle (J X^T J) u, v \rangle = \langle X^* u, v \rangle$$

Wie beim nichtsymmetrischen Standardeigenwertproblem kann eine praktikable hinreichende Bedingung für die vollständige Lösbarkeit des verallgemeinerten Eigenwertproblems nicht gegeben werden, wobei aus den bekannten Gründen die Symmetrie nicht mehr entscheidend ist. (Natürlich gilt, daß im Fall einer nichtsingulären Matrix B n unterschiedliche Nullstellen des charakteristischen Polynoms zwangsläufig zu n linear unabhängigen Eigenvektoren führen müssen, dies läßt sich jedoch nicht direkt aus den Matrizen A und B ablesen!)

Offensichtlich gilt aber der folgende einfache und wichtige Sachverhalt:

Lemma 2.6: Notwendig für die vollständige Lösung des verallgemeinerten Eigenwertproblems ist das Erfülltsein der folgenden Bedingung:

$$\text{Bild}(A) \subseteq \text{Bild}(B)$$

Beweis:

Aus $A = B X \Lambda X^{-1}$ folgt

$$\text{Bild}(A) = \text{Bild}(B X \Lambda X^{-1}) = \{ B X \Lambda X^{-1} x \mid x \in \mathbb{R}^n \} \subseteq \{ B y \mid y \in \mathbb{R}^n \}$$

$$= \text{Bild}(B). \quad \text{Q.E.D.}$$

Satz 2.9 ist insofern allgemeiner als Satz 2.8, als wir für nichtsinguläres B gemäß den Überlegungen nach Lemma 2.1 auf ein symmetrisches verallgemeinertes Eigenwertproblem transformieren können.

Wenn weder die Voraussetzungen von Satz 2.8 (Nichtsingularität von B bzw. schwächer: Nichtsingularität des Büschels (A, B)) noch die von Satz 2.9 (Symmetrie von A und B) gegeben sind, dann kann es sehr wohl passieren, daß ein Paar (A, B) nicht simultan diagonalisierbar ist, obwohl das zugehörige Eigenwertproblem vollständig lösbar ist:

Gegenbeispiel:

Es sei

$$A := \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \quad B := \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},$$

also das Paar (A, B) ein singuläres Büschel.

Dann ist das Eigenwertproblem $A x = \lambda B x$ vollständig lösbar, denn für

$$X := \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}, \quad \Lambda := \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

gilt

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

und

$$B X \Lambda = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},$$

also $A X = B X \Lambda$.

Trotzdem ist (A, B) nicht simultan diagonalisierbar:

Um dies zu beweisen, nehmen wir an, es gäbe zwei nichtsinguläre Matrizen

$$F := \begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix} \quad \text{und} \quad G := \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix}$$

mit der Eigenschaft, daß $F A G$ und $F B G$ diagonal sind.

Aus

$$\begin{aligned} F B G &= \begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} = \begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix} \begin{bmatrix} g_{11} & g_{12} \\ 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} f_{11} g_{11} & f_{11} g_{12} \\ f_{21} g_{11} & f_{21} g_{12} \end{bmatrix} \stackrel{!}{=} \begin{bmatrix} \beta_1 & 0 \\ 0 & \beta_2 \end{bmatrix} \end{aligned}$$

liest man die folgenden Bedingungen ab:

$$(1) \quad f_{11} g_{12} = 0 \qquad (2) \quad f_{21} g_{11} = 0$$

Aus

$$\begin{aligned} F A G &= \begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} = \begin{bmatrix} f_{11} & f_{11} \\ f_{21} & f_{21} \end{bmatrix} \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \\ &= \begin{bmatrix} f_{11}(g_{11} + g_{21}) & f_{11}(g_{12} + g_{22}) \\ f_{21}(g_{11} + g_{21}) & f_{21}(g_{12} + g_{22}) \end{bmatrix} \stackrel{!}{=} \begin{bmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{bmatrix} \end{aligned}$$

ergeben sich die weiteren Bedingungen:

$$(3) \quad f_{11}(g_{12} + g_{22}) = 0 \qquad (4) \quad f_{21}(g_{11} + g_{21}) = 0$$

Es sei zunächst $f_{11} \neq 0$. Dann gilt nach (1), daß $g_{12} = 0$ und damit nach (3) $f_{11} g_{22} = 0$, also $g_{22} = 0$. Dann hat G aber die Gestalt

$$G = \begin{bmatrix} * & 0 \\ * & 0 \end{bmatrix} \quad \text{und ist singulär.}$$

Es bleibt also der Fall $f_{11} = 0$ übrig. $f_{21} = 0$ würde zur Singularität von F führen, also muß $f_{21} \neq 0$ sein, was nach (2) $g_{11} = 0$ impliziert und nach (4) $g_{21} = 0$. Dann wäre G wiederum singulär. Q.E.D.

Die Frage, wann aus der simultanen Diagonalisierbarkeit die vollständige Lösung des Eigenwertproblems folgt, läßt sich restlos beantworten:

Satz 2.10: *Es seien A, B simultan diagonalisierbar.*

Das verallgemeinerte Eigenwertproblem $A x = \lambda B x$ ist genau dann vollständig lösbar, wenn $\text{Bild}(A) \subseteq \text{Bild}(B)$ gilt.

Beweis:

Die Notwendigkeit der Bedingung $\text{Bild}(A) \subseteq \text{Bild}(B)$ ist bereits in Lemma (2.6) festgestellt worden. Wir zeigen jetzt, daß sie hinreichend ist:

Es seien F, G die behaupteten Transformationsmatrizen. Dann gilt

$$F A G = \text{diag}(a_1, \dots, a_n), \quad F B G = \text{diag}(b_1, \dots, b_n).$$

Aus der Voraussetzung $\text{Bild}(A) \subseteq \text{Bild}(B)$ folgt $\text{Bild}(F A G) \subseteq \text{Bild}(F B G)$; dies bedeutet aber, daß das Diagonalelement a_i verschwinden muß, wenn das entsprechende b_i null ist.

Wir setzen jetzt

$$(2.10) \quad \lambda_i = \begin{cases} \frac{a_i}{b_i} & \text{falls } b_i \neq 0 \\ \text{beliebig} & \text{sonst} \end{cases}$$

Damit gilt in jedem Fall $a_i = \lambda_i b_i$.