Eigenvektorzerlegung und selbstadjungierte Abbildungen

Inhaltsverzeichnis

1. Einleitung

In diesem Bericht soll gezeigt werden, daß es zwischen der Eigenvektorzerlegung aus dem Bereich der Neuronalen Netze und den selbstadjungierten Abbildungen der Linearen Algebra einige interessante Verbindungen gibt.

Dabei soll mit Hilfe der *linearen Abbildungen* gezeigt werden, was es bedeutet, wenn ein Neuronales Netz mit einem orthonormierten Eigenvektorsystem als Gewichten, die Ausgabe $\underline{y} = W\underline{x}$ berechnet. Es soll hier versucht werden zu erklären, warum eine gewisse Problematik in dieser Berechnung steckt und warum diese Berechnung trotzdem korrekt ist.

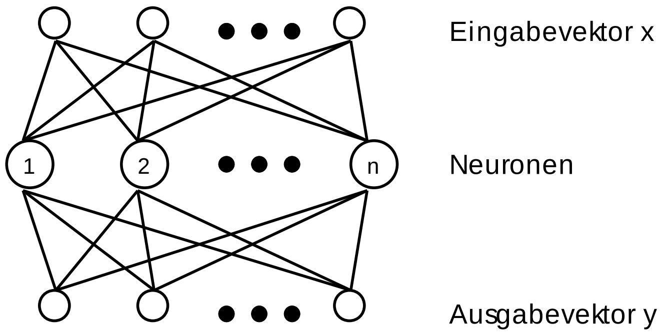
Weiterhin soll versucht werden zu zeigen, welche Rolle das orthonormierte Eigenvektorsystem in der Linearen Algebra spielt und wie einfach es ist, mit Hilfe von selbstadjungierten Abbildungen und symmetrischen Bilinearformen eine Funktion q: $\Re^n \to \Re$ zu definieren, welche als Extrema genau die Eigenvektoren einer symmetrischen Matrix besitzt.

2. Eigenvektorzerlegung

2.1 Leistungen

Es soll hier ein abstraktes neuronales System *N* angenommen werden, welches eine vollständige Eigenvektorzerlegung durchführt, das heißt es soll nur der Fall betrachtet werden, in dem *n* Neuronen *n* Eigenvektoren einer symmetrischen $n \times n$ Matrix C finden. Das C die Korrelationsmatrix der Eingabe ist (siehe *2.2 Berechnungen* weiter unten), ist hier belanglos; wichtig ist nur, daß sie symmetrisch ist.

Die *n* gefundenen Eigenvektoren $\underline{w}_i$ der symmetrischen Matrix C sind normiert und orthogonal und bilden daher eine orthonormierte Basis eines n-dimensionalen Vektorraums V über $\mathfrak{R}$.



Neuronales System

2.2 Berechnungen

Nachdem das neuronale System *N* die *n* Eigenvektoren der symmetrischen Matrix C gefunden hat, werden diese Eigenvektoren ($\underline{w}_1$, $\underline{w}_2$, …,$\underline{w}_n$) dann zeilenweise in die Matrix W geschrieben. Das System *N* berechnet dann aus einer Eingabe $\underline{x}$ eine Ausgabe $\underline{y}$ durch folgende Vorschrift: $\underline{y} = W\underline{x}$.

Sei I die Menge aller Eingaben für das neuronale System *N*. Für I gilt: $I = \{x \in \mathfrak{R}^n\} \subset \mathfrak{R}^n$. Weiterhin sei O die Menge aller Ausgaben. Auch für die Menge O gilt: $O = \{x \in \mathfrak{R}^n\} \subset \mathfrak{R}^n$.

Sei nun $\underline{x} \in I$ und $\underline{y} \in O$ und $\underline{y} = (y_1, y_2, \ldots, y_n) = W\underline{x}$. Dann sind $(y_1, y_2, \ldots, y_n)$ die Koordinaten des Vektors x bezüglich der orthornormierten Basis $W = (\underline{w}_1, \underline{w}_2, \ldots, \underline{w}_n)$[1].

Es gelten weiterhin folgende Eigenschaften:

- Die Kovarianzmatrix C berechnet sich durch: $C_{ij} = \langle (x_i - \langle x_i \rangle)\,(x_j - \langle x_j \rangle) \rangle_x$, wobei sich mit $\langle x_i \rangle = 0$ für $i = 1, \ldots, n$ die Korrelationsmatrix C berechnet durch $C_{ij} = \langle x_i\, x_j \rangle_x$.

[1] Diese Aussage wird im Kapitel *Lineare Algebra* weiter untersucht und auch bewiesen.

- Die Ausgabe y besitzt maximale Varianz, also $\langle(y - \langle y\rangle)^2\rangle_y = max$ für alle w mit $|w|^2 = 1$.
- Die Komponenten der y's sind unkorreliert, also $\langle(y_i - \langle y_i\rangle)(y_j - \langle y_j\rangle)\rangle_y = 0$ für $i \neq j$.

3. Lineare Algebra

Um die Notation zu vereinfachen, werden in diesem Kapitel alle Vektoren mit griechischen und alle Skalare mit römischen Buchstaben angegeben. Es wird vorausgesetzt, daß die Begriffe *Vektorraum*, *Lineare Unabhängigkeit* und *Basis* bekannt sind.

3.1 Dimension

Definition: Die Anzahl der Elemente in einer Basis von V heißt *Dimension* von V, dim V = n.

Korollar: Es sei V ein Vektorraum der Dimension n. Dann gelten folgende Aussagen:
(i) Weniger als n Vektoren bilden kein Erzeugendensystem[2] .
(ii) Mehr als n Vektoren sind linear abhängig.
(iii) Jedes Erzeugendensystem von n Vektoren ist eine Basis von V.
(iv) Jede linear unabhängige Familie von n Vektoren ist eine Basis von V.

Korollar: Es sei W ein Unterraum[3] des Vektorraums V, dim V = n. Dann gilt dim W $\leq$ dim V. Gleichheit gilt genau dann, wenn W = V.

3.2 Isomorphismen und Koordinaten

Definition: Es sei V ein Vektorraum über F, und es sei $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ eine (geordnete) Basis. Jeder Vektor $\alpha \in V$ läßt sich in eindeutiger Weise als Linearkombination von $\alpha_1, \alpha_2, \ldots, \alpha_n$ darstellen[4]: $\alpha = \sum_i a_i \alpha_i$, $a_i \in F$. Die Koeffizienten $a_1, a_2, \ldots, a_n$ heißen *Koordinaten* des Vektors α bezüglich der Basis $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$.

Definition: Es seinen V, W zwei Vektorräume über dem Körper F. Eine Abbildung $f: V \to W$ heißt *linear*, wenn folgende zwei Bedingungen erfüllt sind:
(i) $f(\xi + \eta) = f(\xi) + f(\eta)$, $\xi, \eta \in V$.
(ii) $f(c\, \xi) = c\, f(\xi)$, $c \in F$, $\xi \in V$.

Satz: Ist $f: V \to W$ eine lineare Abbildung, so führt f den Nullvektor in V in den Nullvektor von W über; f(0) = 0. Linearkombinationen in V werden in Linearkombinationen in W überführt:
$f(\sum_i a_i \alpha_i) = \sum_i a_i f(\alpha_i)$, $\alpha_i \in V$, $a_i \in F$.

Definition: Die *Identität* $1_V: V \to V$ ist definiert durch $1_V(\alpha) = \alpha$ für alle $\alpha \in V$.

Definition[5]: Eine lineare Abbildung $f: V \to W$ heißt ein *Isomorphismus* , wenn es eine (lineare) Abbildung $g: W \to V$ gibt, mit $g \circ f = 1_V$ und $f \circ g\ 1_W$.

[2] Eine Familie S von Vektoren ist genau dann ein Erzeugendensystem eines Vektorraums V, falls S den Raum [S] aufspannt und es gilt: [S] = V.
[3] Eine nichtleere Untermenge W eines Vektorraums V über einem Körper F heißt Unterraum von V, wenn für $\alpha, \beta \in W$ und $a, b \in F$ stets gilt: $a\alpha + b\beta \in W$.
[4] Diese Aussage müßte korrekterweise bewiesen werden.
[5] Es kann bewiesen werden, daß g dann auch linear und ein Isomorphismus ist.

Bemerkung: Es kann bewiesen werden, daß unter einem Isomorphismus eine linear unabhängige Familie von Vektoren in eine linear unabhängige Familie übergeht. Dies gilt im allgemeinen nicht für lineare Abbildungen.

3.3 Zeilen- und Spaltenrang einer Matrix

Definition: Es sei A eine $m \times n$ Matrix über F. Es sei S die Familie der Zeilen, aufgefasst als Elemente in F^n: $S = \{\alpha_1, \alpha_2, \ldots, \alpha_m\}$, $\alpha_i = (a_{i1}, a_{i2}, \ldots, a_{in})$, $1 \leq i \leq m$. Es sei T die Familie der Spalten, aufgefasst als Elemente in F^m: $T = \{\beta_1, \beta_2, \ldots, \beta_n\}$, $\beta_k = (a_{1k}, a_{2k}, \ldots, a_{mk})$, $1 \leq k \leq n$. Dann heißt [S] der Zeilenraum und dim [S] der *Zeilenrang* von A. Analog heißt [T] der Spaltenraum und dim [T] der *Spaltenrang* von A.

Satz: Für jede Matrix A ist der Zeilenrang gleich dem Spaltenrang.

Definition: Der gemeinsame Wert von Zeilenrang und Spaltenrang einer Matrix A heißt *Rang* der Matrix, Rang A. Eine $n \times n$ Matrix A mit Rang A = n heißt *regulär*, sonst heißt sie *singulär*.

Satz: Für eine $n \times n$ Matrix A sind folgende Aussagen äquivalent:
(i) A ist regulär.
(ii) Rang A = n.
(iii) Die Zeilen sind linear unabhängig.
(iv) Die Spalten sind linear unabhängig.
(v) Das lineare Gleichungssystem $Ax = 0$ hat nur die triviale Lösung.

3.4 Basiswechsel

Bemerkung: Es sei V ein Vektorraum über F und es seien S und T zwei Basen von V, $S = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$, $T = \{\beta_1, \beta_2, \ldots, \beta_n\}$. Jeder Vektor $\xi \in V$ läßt sich dann sowohl als Linearkombination von Vektoren aus S als auch als Linearkombination von Vektoren aus T darstellen: $\xi = \sum_k x_k \alpha_k = \sum_i y_i \beta_i$, wobei beide Darstellungen eindeutig sind. Die Koordinaten von ξ bezüglich S sind $x_1, x_2, \ldots, x_n$, diejenigen bezüglich T sind $y_1, y_2, \ldots, y_n$.

Da T eine Basis ist, läßt sich α_k , $1 \leq k \leq n$, als Linearkombination der Vektoren $\beta_1, \beta_2, \ldots, \beta_n$ darstellen: $\alpha_k = \sum_i a_{ik} \beta_i$. Die alte Basis S wird durch die neue Basis T ausgedrückt. Zum *Basiswechsel* von S nach T gehört die $n \times n$ Matrix $[a_{ik}]$.

Wir erhalten nun: $\xi = \sum_k x_k \alpha_k = \sum_k x_k \sum_i a_{ik} \beta_i = \sum_i (\sum_k a_{ik} x_k) \beta_i = \sum_i y_i \beta_i$, mit $y_i = \sum_k a_{ik} x_k$. Die neuen Koordinaten werden durch die alten Koordinaten mit Hilfe der Matrix $[a_{ik}]$ ausgedrückt, die zum Basiswechsel von S nach T gehört.

Definition: Es sein A eine $n \times n$ Matrix. Eine $n \times n$ Matrix B mit der Eigenschaft $BA = I = AB$ heißt die zu A *inverse Matrix*; sie wird mit A^{-1} bezeichnet.

Satz: Für eine $n \times n$ Matrix A über F gilt die Äquivalenz folgender Aussagen:
(i) A ist regulär.
(ii) Zu A existiert eine eindeutig bestimmte $n \times n$ Matrix B mit $BA = I = AB$.
(iii) A ist Matrix eines Basiswechsels in einem n-dimensionalen Vektorraum über F.

3.5 Lineare Abbildungen

Definition: Es sei f: $V \to W$ eine lineare Abbildung. Man definiert den *Kern* von f durch
$\ker f = \{\xi \in V \mid f(\xi) = 0\}$
und das *Bild* von f durch
$\operatorname{im} f = \{f(\xi) \in W \mid \xi \in V\} = f(V)$.

Satz: Ist f: $V \to W$ eine lineare Abbildung, so ist ker f ein Unterraum von V und im f ein Unterraum von W.

Definition: Die (lineare) Abbildung f: $V \to W$ heißt *injektiv*, wenn aus $f(\alpha) = f(\beta)$ stets $\alpha = \beta$ folgt, $\alpha, \beta \in V$.

Satz: Die lineare Abbildung f: $V \to W$ ist genau dann injektiv, wenn ker f = 0.

Definition: Die (lineare) Abbildung f: $V \to W$ heißt *surjektiv*, wenn zu jedem $\eta \in W$ mindestens ein $\xi \in V$ existiert, mit $f(\xi) = \eta$.

Satz: Es sei f: $V \to W$ eine lineare Abbildung. Genau dann ist f ein Isomorphismus, wenn f injektiv und surjektiv ist.

Satz[6]: Es sei $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ eine Basis von V und $\beta_1, \beta_2, \ldots, \beta_n$ beliebige Vektoren von W. Dann gibt es eine eindeutig bestimmte lineare Abbildung f: $V \to W$ mit $f(\alpha_i) = \beta_i$, i = 1, …, n.

Satz[7]: Es sei f: $V \to W$ eine lineare Abbildung und dim V = n. Dann gilt:
dim V = n = dim (ker f) + dim (im f).

3.6 Matrizen

Bemerkung[8]: Sei f: $V \to W$ eine lineare Abbildung , und es seien $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ eine Basis von V und $\{\beta_1, \beta_2, \ldots, \beta_m\}$ eine Basis von W. Nach dem vorletzten Satz aus Abschnitt 3.5 ist f eindeutig bestimmt durch die Bilder $f(\alpha_k)$ der Basisvektoren. Es gelte:
$f(\alpha_k) = \sum_i^m a_{ik}\, \beta_i$, k = 1,…, n.

Ist α ein beliebiger Vektor aus V, so läßt sich $f(\alpha)$ wie folgt bestimmen. Es sei
$\alpha = \sum_k^n x_k\, \alpha_k$, $f(\alpha) = \sum_i^m y_i\, \beta_i$. Dann gilt:

$$f(\alpha) = f(\textstyle\sum_k^n x_k\, \alpha_k) = \sum_k^n x_k\, f(\alpha_k) = \sum_k^n x_k \sum_i^m a_{ik}\, \beta_i = \sum_i^m (\sum_k^n a_{ik}\, x_k)\, \beta_i.$$

Für die Koordinaten $y_1, \ldots, y_m$ von $f(\alpha)$ bezüglich der Basis $\{\beta_1, \beta_2, \ldots, \beta_m\}$ erhalten wir somit: $y_i = \sum_k^n a_{ik}\, x_k$, i = 1, …, m.

6 Dieser Satz enthält eine sehr wichtige Aussage: die Abbildung f wird eindeutig durch die Bilder der Basisvektoren beschrieben. Kennt man die Bilder der Basisvektoren, so kennt man auch f.

7 Der Bildraum einer lineare Abbildung f verliert genau die Dimensionen, die im Kern der Abbildung liegen. Oder anderst: die Abbildung einiger Vektoren auf 0 reduziert die Dimension des Bildraums.

8 Diese Bemerkung ist sehr wichtig und wird in Kapital 4.1 verwendet.

Satz: Ist $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ eine Basis von V und $\{\beta_1, \beta_2, \ldots, \beta_m\}$ eine Basis von W, dann wird die lineare Abbildung f: $V \to W$ vollständig durch eine $m \times n$ Matrix A beschrieben. Dabei stehen in der k-ten Spalte von A gerade die Koordinaten des Vektors $f(\alpha_k)$ bezüglich der Basis $\{\beta_1, \beta_2, \ldots, \beta_m\}$.

Definition: Die Matrix A heißt die *zu f gehörige Matrix* bezüglich der Basen $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ und $\{\beta_1, \beta_2, \ldots, \beta_m\}$.

Bemerkung: Schreiben wir die Koordinaten des Vektors α als $n \times 1$ Matrix und die Koordinaten des Vektors $f(\alpha)$ als $m \times 1$ Matrix, so wird die Abbildung f beschrieben durch $Ax = y$.

Satz: Ist S eine Basis von V und T eine Basis von W, so läßt sich zu jeder $m \times n$ Matrix A eine lineare Abbildung f: $V \to W$ so definieren, daß A gerade die zu f gehörige Matrix bezüglich S und T ist.

3.7 Rang einer linearen Abbildung

Definition: Für eine lineare Abbildung f definiert man den *Rang* durch Rang f = dim (im f).

Satz: Zu der linearen Abbildung f: $V \to W$ gehöre bezüglich gegebener Basen S und T die Matrix A. Dann gilt: Rang f = Rang A.

Satz: Es gilt: Rang f + dim(ker f) = dim V.

Satz: Es sei f: $V \to W$ eine lineare Abbildung; es sei dim V = dim W = n. Dann sind die folgenden Aussagen äquivalent:

(i) f ist ein Isomorphismus.
(ii) f ist injektiv.
(iii) f ist surjektiv.

3.8 Euklidische Vektorräume

Definition: Es sei V ein Vektorraum über $\Re$. Ein (reelles) *Skalarprodukt* in V ist eine Funktion (,): $V \times V \to \Re$, welche die folgenden Eigenschaften besitzt:

(i)	$(\alpha, \beta) = (\beta, \alpha)$	$\alpha, \beta \in V$ (symmetrisch)
(ii)	$(a\alpha + b\beta, \gamma) = a\,(\alpha, \gamma) + b\,(\beta, \gamma)$	$\alpha, \beta, \gamma \in V$, $a, b \in \Re$ (linear)
(iii)	$(\alpha, \alpha) \geq 0$; $(\alpha, \alpha) = 0 \Leftrightarrow \alpha = 0$	$\alpha \in V$ (positiv definit)

Ein endlich-dimensionaler Vektorraum V über $\Re$ zusammen mit einem Skalarprodukt heißt *euklidischer* Vektorraum.

Bemerkung: Aus den Eigenschaften (i), (ii) eines reellen Skalarprodukts folgt für $\alpha, \beta, \gamma \in V$ und $a, b \in \Re$ sofort $(\gamma, a\alpha + b\beta) = a\,(\gamma, \alpha) + b\,(\gamma, \beta)$. Das Skalarprodukt ist somit linear in beiden Argumenten; man sagt auch, es sei bilinear. Man beweist nun ohne Schwierigkeit, daß $(\sum_i^m a_i\,\alpha_i, \sum_j^n b_j\,\beta_j) = \sum_i^m \sum_j^n a_i\, b_j\,(\alpha_i, \beta_j)$ gilt, wobei $\sum_i^m a_i\,\alpha_i$ und $\sum_j^n b_j\,\beta_j$ beliebige Linearkombinationen von Vektoren aus V sind. Aus dieser Beziehung folgt insbesondere, daß ein Skalarprodukt in V bestimmt ist, wenn seine Werte auf einer Basis gegeben sind.

Definition: Es sei (,): $V \times V \to \Re$ ein Skalarprodukt und $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ eine Basis von V. Die Matrix $[c_{ij}]$, definiert durch $c_{ij} = (\alpha_i, \alpha_j)$, $1 \le i, j \le n$, heißt *Matrix C des Skalarprodukts* bezüglich der Basis $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$.

Satz: Die Matrix eines reellen Skalarprodukts ist symmetrisch.

3.9 Norm

Satz: Es sei V ein Vektorraum mit einem Skalarprodukt. Dann gilt $|(\alpha, \beta)|^2 \le (\alpha, \alpha)\,(\beta, \beta)$. Die Gleichheit tritt genau dann ein, wenn α und β linear abhängig sind (Cauchy-Schwarz-Ungleichung).

Definition: Es sei V ein reeller oder komplexer Vektorraum. Eine Funktion $\| \; \|: V \to \Re$ heißt *Norm*, wenn folgende Bedingungen erfüllt sind:
(i) $\|a\alpha\| = |a|\ \|\alpha\|$
(ii) $\|\alpha + \beta\| \le \|\alpha\| + \|\beta\|$ (Dreiecksungleichung)
(iii) $\|\alpha\| = 0 \Leftrightarrow \alpha = 0$

Satz: Es sei V ein Vektorraum mit Skalarprodukt (,). Dann ist die Funktion $\| \; \|: V \to \Re$, definiert durch $\|\alpha\| = \sqrt{(\alpha,\alpha)}$, eine Norm.

3.10 Orthonormierte Basen

Definition: Es sei V ein Vektorraum mit Skalarprodukt (,). Der Vektor $\alpha \in V$ heißt *orthogonal* zum Vektor $\beta \in V$, wenn $(\alpha, \beta) = 0$. Die Teilmenge S von V heißt *orthogonal* zur Teilmenge T von V, wenn $(\alpha, \beta) = 0$ für alle $\alpha \in S$ und $\beta \in T$.

Definition: Eine nichtleere Familie S von V heißt orthogonal, wenn je zwei verschiedene Elemente α, β der Familie S orthogonal sind. Eine orthogonale Familie S von Vektoren aus V heißt *orthonormiert*, wenn $(\alpha, \alpha) = 1$ für alle $\alpha \in S$. Eine orthonormierte Familie S von Vektoren aus V, die zugleich Basis von V ist, heißt eine *orthonormierte Basis* von V.

Bemerkung: Eine orthogonale Familie S von Vektoren aus V, die den Nullvektor nicht enthält, kann in eine orthonormierte Familie S' überführt werden, indem man $\alpha' = \alpha / \|\alpha\|$, $\alpha \in S$, $\alpha' \in S'$, setzt. Man sagt S' sei durch Normieren aus S hervorgegangen.

Satz: Eine orthogonale Familie S von Vektoren aus V, die den Nullvektor nicht enthält, ist linear unabhängig.

Beweis: Es sei $\sum_i a_i \alpha_i = 0$, $\alpha_i \in S$. Dann gilt für $j = 1, \ldots, n$:
$0 = (0, \alpha_j) = (\sum_i a_i \alpha_i, \alpha_j) = \sum_i a_i (\alpha_i, \alpha_j) = a_j (\alpha_j, \alpha_j)$. Da $\alpha_j \neq 0$, folgt $a_j = 0$.

Satz[9]: Es sei V ein endlich-dimensionaler Vektorraum mit Skalarprodukt, und es sei $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ eine orthonormierte Basis. Dann gilt für jedes $\beta \in V$: $\beta = \sum_i^n (\beta, \alpha_i)\, \alpha_i$.

[9] Dieser Satz ist sehr wichtig und wird in Kapitel 4.1 verwendet.

Beweis: Es sei $\beta = \sum_j^n xj\ \alpha_j$.
Dann gilt für i = 1, …, n: $(\beta, \alpha_i) = (\sum_j^n x_j\ \alpha_j, \alpha_i) = \sum_j^n x_j\ (\alpha_j, \alpha_i) = x_i$.

Theorem: Jeder endlich-dimensionale Vektorraum mit Skalarprodukt besitzt eine orthonormierte Basis.

Beweis[10]: Der Beweis erfolgt mit Hilfe des Gram-Schmidt'schen Orthogonalisierungsverfahren. Eine so konstruierte orthogonale Basis wird dann anschließend normiert.

Satz: Es sei V ein endlich-dimensionaler Vektorraum mit Skalarprodukt, und es sei $\{\beta_1, \beta_2, \ldots, \beta_n\}$ eine orthonormierte Basis. Die Vektoren $\alpha_1, \alpha_2, \ldots, \alpha_n$ seien gegeben durch $\alpha_k = \sum_i^n a_{ik}\ \beta_i$, $1 \leq k \leq n$. Es sei $A = [a_{ik}]$. Die Familie $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ ist genau dann eine orthonormierte Basis von V, wenn folgendes gilt: $A^T A = I$.

Definition: Es sei A eine $n \times n$ Matrix mit $A^T = A^{-1}$. Ist A eine reelle Matrix, so heißt A *orthogonal*; ist A eine komplexe Matrix, so heißt A *unitär*.

Satz: Es sei A eine reelle $n \times n$ Matrix. Dann sind folgende Aussagen äquivalent:
(i) A ist orthogonal.
(ii) $A^T A = I$.
(iii) $AA^T = I$.
(iv) A vermittelt einen Basiswechsel zwischen orthonormierten Basen eines n-dimensionalen euklidischen Vektorraums.
(v) Die Spalten der Matrix A bilden eine orthonormierte Basis des Vektorraums $\Re^n$, versehen mit dem Standard-Skalarprodukt.
(vi) Die Zeilen der Matrix A bilden eine orthonormierte Basis des Vektorraums $\Re^n$, versehen mit dem Standard-Skalarprodukt.

3.11 Orthogonale Selbstabbildungen

Definition: Es sei V ein Vektorraum mit Skalarprodukt. Eine lineare Selbstabbildung $f: V \to V$ heißt *orthogonal*, wenn $(f(\alpha), f(\beta)) = (\alpha, \beta)$ für alle $\alpha, \beta \in V$.

Satz: Es sei V ein endlich-dimensionaler Vektorraum mit Skalarprodukt, und es sei $f: V \to V$ eine orthogonale Selbstabbildung. Dann gilt:
(i) $\|f(\alpha)\| = \|\alpha\|$.
(ii) Sind α, β orthogonal, so sind auch $f(\alpha), f(\beta)$ orthogonal.
(iii) Ist λ ein Eigenwert von f, dann ist $|\lambda| = 1$.
(iv) f ist ein Isomorphismus.

Satz: Es sei $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ eine orthonormierte Basis von V. Genau dann ist $f: V \to V$ orthogonal, wenn gilt: $(f(\alpha_k), f(\alpha_l)) = \delta_{kl}$.

Satz: Es sei V ein endlich-dimensionaler Vektorraum mit Skalarprodukt, und es sei S eine orthonormierte Basis. Die lineare Selbstabbildung $f: V \to V$ ist genau dann orthogonal, wenn die Matrix A von f bezüglich S orthogonal ist, das heißt wenn gilt: $A^T A = I$.

[10] Das Gram-Schmid'sche Orthogonalisierungsverfahren wird als bekannt vorausgesetzt.

Satz: Ist die Matrix A orthogonal, so gilt: $|\det A| = 1$.

Beweis: $1 = \det I = \det(A^T A) = \det A^T \det A \Rightarrow |\det A| = 1$.

3.12 Adjungierte Selbstabbildungen

Definition: Die zu $f: V \to V$ *adjungierte Abbildung* $f^*: V \to V$ ist definiert durch die Beziehung $(f(\xi), \eta) = (\xi, f^*(\eta))$, $\xi, \eta \in V$.

Satz: Es sei V ein endlich-dimensionaler Vektorraum mit Skalarprodukt, und es sei S eine orthonormierte Basis. Wird $f: V \to V$ bezüglich S durch die Matrix A dargestellt, so wird die adjungierte Abbildung $f^*: V \to V$ bezüglich S durch die Matrix A^T dargestellt.

Beweis: Es sei $S = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$, $A = [a_{ik}]$ und $B = [b_{jl}]$; es gilt dann also:
$f(\alpha_k) = \sum_i^n a_{ik}\, \alpha_i$, $1 \le k \le n$.

Ferner gelte $f^*(\alpha_l) = \sum_j^n b_{jl}\, \alpha_j$, $1 \le l \le n$. Damit erhält man:

$(f(\alpha_k), \alpha_l) = (\sum_i^n a_{ik}\, \alpha_i, \alpha_l) = \sum_i^n a_{ik}\, (\alpha_i, \alpha_l) = a_{lk}$,

$(\alpha_k, f^*(\alpha_l)) = (\alpha_k, \sum_j^n b_{jl}\, \alpha_j) = \sum_j^n b_{jl}\, (\alpha_k, \alpha_j) = b_{kl}$.

Wegen $(f(\alpha_k), \alpha_l) = (\alpha_k, f^*(\alpha_l))$ ergibt sich daraus $a_{lk} = b_{kl}$. Die Abbildung f^* wird somit bezüglich S durch A^T dargestellt. Damit ist der Satz bewiesen.

Satz[11]: Es sei V ein endlich-dimensionaler Vektorraum mit Skalarprodukt, und es sei $f: V \to V$ eine lineare Selbstabbildung. Dann haben f und f^* den gleichen Rang.

3.13 Selbstadjungierte Selbstabbildungen

Definition: Es sei V ein Vektorraum mit Skalarprodukt. Die lineare Selbstabbildung $f: V \to V$ heißt *selbstadjungiert*, wenn für alle $\xi, \eta \in V$ gilt: $(f(\xi), \eta) = (\xi, f(\eta))$, das heißt wenn $f = f^*$.

Satz[12]: Es sei V ein endlich-dimensionaler Vektorraum mit Skalarprodukt und S eine orthonormierte Basis. Die lineare Selbstabbildung $f: V \to V$ werde bezüglich S durch die Matrix A beschrieben. Genau dann ist $f: V \to V$ selbstadjungiert, wenn gilt $A = A^T$.

Beweis: Es sei $S = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$, $A = [a_{ik}]$.

$(f(\alpha_k), \alpha_j) = (\sum_i^n a_{ik}\, \alpha_i, \alpha_j) = \sum_i^n a_{ik}\, \delta_{ij} = a_{jk}$,

$(\alpha_k, f(\alpha_j)) = (\alpha_k, \sum_i^n a_{ij}\, \alpha_i) = \sum_i^n a_{ij}\, \delta_{ki} = a_{kj}$.

[11] Da der Beweis die *orthogonalen Komplemente* benötigt, wird an dieser Stelle darauf verzichtet.
[12] Eine Matrix A, für die gilt $A = A^T$ ist symmetrisch.

Daraus folgt, daß $a_{jk} = a_{kj}$, falls f selbstadjungiert ist. Gilt nun umgekehrt $A = A^T$, so folgt aus obigem: $(f(\alpha_k), \alpha_j) = (\alpha_k, f(\alpha_j))$, $1 \le k, j \le n$. Nun muß noch gezeigt werden, daß das für alle ξ, $\eta \in V$ gilt. Sei nun also $\xi = \sum_k^n x_k \alpha_k$ und $\eta = \sum_j^n y_j \alpha_j$. Dann gilt:

$(f(\xi), \eta) = (f(\sum_k^n x_k \alpha_k), \sum_j^n y_j \alpha_j) = \sum_k^n x_k \sum_j^n y_j (f(\alpha_k), \alpha_j) = \sum_k^n x_k \sum_j^n y_j (\alpha_k, f(\alpha_j)) =$

$(\sum_k^n x_k \alpha_k, f(\sum_j^n y_j \alpha_j)) = (\xi, f(\eta))$.

Da $(f(\xi), \eta) = (\xi, f(\eta))$ für alle $\xi, \eta \in V$ gilt, folgt, daß f selbstadjungiert ist.

Satz: Es sei V ein Vektorraum der Dimension n mit Skalarprodukt, und es sei $f: V \to V$ eine selbstadjungierte Selbstabbildung. Dann sind alle Nullstellen des charakteristischen Polynoms $p_f(x)$ reell; es gilt dann also[13]: $p_f(x) = \pm (x - \lambda_1) \cdots (x - \lambda_n)$ mit $\lambda_1 \ldots \lambda_n \in \Re$.

Theorem: Es sei V ein endlich-dimensionaler Vektorraum mit Skalarprodukt, und es sei f eine selbstadjungierte Selbstabbildung. Dann existiert eine orthonormierte Basis von V, die aus den Eigenvektoren von f besteht.

Beweis: Es wird mit Induktion nach der Dimension n des Vektorraums V bewiesen, daß eine Basis der verlangten Art existiert. Ist n = 1, so gibt es nichts zu beweisen. Für $n \ge 2$ läßt sich das charakteristische Polynom in Linearfaktoren aufspalten:
$p_f(x) = \pm (x - \lambda_1) \cdots (x - \lambda_n)$ mit $\lambda_1 \ldots \lambda_n \in \Re$.

Es sei nun α_1 mit $\|\alpha_1\| = 1$ ein Eigenvektor zu λ_1 und $W = \{\xi \in V \mid (\alpha_1, \xi) = 0\}$. Da die Formel $\dim V = \dim U + \dim U^\perp$ gilt[14], ist $\dim W = n - 1$. W ist ein bezüglich f invarianter Teilraum. Ist nämlich $\xi \in W$, so folgt $(\alpha_1, f(\xi)) = (f(\alpha_1), \xi) = (\lambda_1 \alpha_1, \xi) = \lambda_1 (\alpha_1, \xi) = 0$.

Es ist also $f(\xi) \in W$. damit induziert f eine selbstadjungierte Selbstabbildung $f|_W: W \to W$. Nach Induktionsvoraussetzung existiert in W eine orthonormierte Basis $\{\alpha_2, \ldots, \alpha_n\}$, die aus den Eigenvektoren von $f|_W$ besteht. Da jeder Eigenvektor von $f|_W$ auch Eigenvektor von f ist, ist die Familie $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ eine Basis von V der verlangten Art. Damit ist das Theorem bewiesen.

3.14 Bilinearformen; quadratische Formen

Definition: Es sei V ein Vektorraum über $\Re$. Eine *Bilinearform* auf V ist eine Funktion $\phi: V \times V \to \Re$, die jedem (geordneten) Paar von Vektoren aus V eine reelle Zahl zuordnet, so daß für $\alpha, \beta, \gamma \in V$, $a, b, c \in \Re$ folgende Bedingungen erfüllt sind (Bilinearität):
(i) $\phi(a\alpha + b\beta, \gamma) = a\,\phi(\alpha, \gamma) + b\,\phi(\beta, \gamma)$,
(ii) $\phi(\alpha, b\beta + c\gamma) = b\,\phi(\alpha, \beta) + c\,\phi(\alpha, \gamma)$.
Eine Bilinearform heißt symmetrisch, wenn gilt:
(iii) $\phi(\alpha, \beta) = \phi(\beta, \alpha)$.

[13] Nach dem Fundamentalsatz der Algebra läßt sich das komplexe charakteristische Polynom in Linearfaktoren aufspalten. Für das charakteristische Polynom von selbstadjungierte Selbstabbildungen kann eine Aufspaltung in Linearfaktoren immer mit reellen Nullstellen durchgeführt werden.

[14] Mit Hilfe eines Satzes über *orthogonale Komplemente* läßt sich diese Behauptung beweisen.

Bemerkung: Es sei $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ eine Basis von V, und es gelte $\xi = \sum_i^n x_i\, \alpha_i$, $\eta = \sum_j^n y_j\alpha_j$. Ist ϕ: $V \times V \to \Re$ eine Bilinearform auf V, so folgt wegen der Bilinearität $\phi(\xi, \eta) = \sum_i^n \sum_j^n x_i y_j\, \phi(\alpha_i, \alpha_j)$. Die Bilinearform ϕ ist somit durch ihre Werte auf den Basiselemeten bestimmt. Eine $n \times n$ Matrix $[c_{ij}]$, definiert durch $c_{ij} = \phi(\alpha_i, \alpha_j)$, heißt die Matrix der Bilinearform bezüglich der Basis $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$. Umgekehrt gehört zu jeder $n \times n$ Matrix $[c_{ij}]$ eine Bilinearform ϕ auf V, definiert durch $\phi(\xi, \eta) = \sum_i^n \sum_j^n x_i y_j\, c_{ij}$, so daß deren Matrix bezüglich der Basis $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ gerade $[c_{ij}]$ ist.

Bemerkung: Die Bilinearform ϕ auf V ist genau dann symmetrisch, wenn die Matrix von ϕ bezüglich irgendeiner Basis von V symmetrisch ist.

Definition: Es sei ϕ: $V \times V \to \Re$ eine Bilinearform auf V. Die Funktion Q: $V \to \Re$, definiert durch $Q(\xi) = \phi(\xi, \xi)$, $\xi \in V$, heißt die zu ϕ gehörende *quadratische Form*.

3.15 Bilinearformen; euklidische Vektorräume

Satz: Es sei ϕ eine Bilinearform auf dem endlich-dimensionalen (reellen) Vektorraum V. Ist (,) ein Skalarprodukt in V, so gibt es eindeutig bestimmte lineare Selbstabbildungen f, g: $V \to V$ mit $\phi(\xi, \eta) = (\xi, f(\eta)) = (g(\xi), \eta)$, $\xi, \eta \in V$. Ferner ist f (und g) genau dann selbstadjungiert, wenn ϕ symmetrisch ist.

Beweis[15]: Es sei $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ eine orthonormierte Basis von V, und es gelte:
$\phi(\alpha_i, \alpha_k) = a_{ik}$, $1 \le i, k \le n$.
Man definiert f, g: $V \to V$ durch: $f(\alpha_k) = \sum_j^n a_{jk}\, \alpha_j$, $1 \le k \le n$, und $g(\alpha_i) = \sum_j^n a_{ij}\, \alpha_j$, $1 \le i \le n$.
Dann gilt:
$(\alpha i, f(\alpha_k)) = (\alpha i, \sum_j^n a_{jk}\, \alpha_j) = \sum_j^n a_{jk}\, \delta_{ij} = a_{ik} = \phi(\alpha_i, \alpha_k)$,

$(g(\alpha_i), \alpha_k) = (\sum_j^n a_{ij}\, \alpha_j, \alpha_k) = \sum_j^n a_{ij}\, \delta_{jk} = a_{ik} = \phi(\alpha_i, \alpha_k)$.

Wegen der Bilinearität von ϕ folgt sofort $\phi(\xi, \eta) = (\xi, f(\eta)) = (g(\xi), \eta)$ für alle $\xi, \eta \in V$.

Ist ϕ symmetrisch, so gilt für $\xi, \eta \in V$: $(\xi, f(\eta)) = \phi(\xi, \eta) = \phi(\eta, \xi) = (\eta, f(\xi)) = (f(\xi), \eta)$, also ist f selbstadjungiert.

Ist nun f selbstadjungiert, so gilt: $\phi(\xi, \eta) = (\xi, f(\eta)) = (f(\xi), \eta) = (\eta, f(\xi)) = \phi(\eta, \xi)$, also ist ϕ symmetrisch. Damit ist der Satz vollständig bewiesen.

Bemerkung: Es geht aus dem Beweis des obigen Satzes hervor, daß für eine orthonormierte Basis die Matrix von f mit der Matrix von ϕ übereinstimmt, während die Matrix von g zu ihr transponiert ist.

[15] Auf den Beweis der Eindeutigkeit von f und g wurde hier verzichtet.

3.16 Extremaleigenschaft der Eigenwerte

Bemerkung: Mit V soll der Vektorraum $\Re^n$ bezeichnet werden. Dieser wird mit dem Standard-Skalarprodukt (,) versehen. Nach dem Satz aus dem vorherigen Kapitel gibt es dann eine eindeutige Beziehung zwischen den selbstadjungierten Selbstabbildungen f: $V \rightarrow V$ und den symmetrischen Bilinearformen ϕ: $V \times V \rightarrow \Re$. Dabei gilt: $\phi(\xi, \xi) = (\xi, f(\xi))$, $\xi \in V$. Bezüglich der Standardbasis in V werden f und ϕ durch die gleiche reelle symmetrische $n \times n$ Matrix beschrieben und umgekehrt definiert jede solche Matrix eine selbstadjungierte Selbstabbildung f von V und eine symmetrische Bilinearform ϕ auf V, so daß die obige Gleichung erfüllt ist. Es wird in diesem Abschnitt der Quotient q, gegeben durch:

$q(\xi) = \dfrac{\phi(\xi,\xi)}{(\xi,\xi)} = \dfrac{(\xi, f(\xi))}{(\xi,\xi)}$, $\xi \in V$, $\xi \neq 0$ betrachtet und dessen Eigenschaften hergeleitet.

Bemerkung: Ist α ein Eigenvektor von f zum Eigenwert λ, so gilt offensichtlich:

$q(\alpha) = \dfrac{(\alpha, f(\alpha))}{(\alpha,\alpha)} = \dfrac{\lambda(\alpha,\alpha)}{(\alpha,\alpha)} = \lambda$. Die Eigenwerte von f liegen somit im Wertebereich von q.

Bemerkung: Für $c \in \Re$, $c \neq 0$, gilt:

$q(c\xi) = \dfrac{(c\xi, f(c\xi))}{(c\xi, c\xi)} = \dfrac{c^2(\xi, f(\xi))}{c^2(\xi,\xi)} = q(\xi)$.

Um das Verhalten der Funktion q zu studieren, genügt es also, die Vektoren mit Norm 1 zu betrachten. Wir setzen deshalb $M = \{\xi \in V \mid \|\xi\| = 1\}$ und betrachten q im folgenden als reellwertige Funktion auf M.

Bemerkung: Führen wir in $\Re^n$ die übliche Topologie ein, so stellt man leicht fest, daß M kompakt und q stetig ist. Aus einem wohlbekannten Satz der Analysis folgt dann sofort, daß q auf M sowohl ein Minimum als auch ein Maximum annimmt[16].

Satz: Das Maximum (bzw. Minimum) der Funktion q auf M ist der größte (bzw. kleinste) Eigenwert von f. Die Maximalstellen (bzw. Minimalstellen) sind die zugehörigen Eigenvektoren.

Satz: Sei $V = \Re^n$ ein Vektorraum mit Standardbasis $I = \{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n\}$ und Standard-Skalarprodukt (,): $\Re^n \times \Re^n \rightarrow \Re$. Bezüglich I sei f: $\Re^n \rightarrow \Re^n$ eine selbstadjungierte Selbstabbildung und ϕ: $\Re^n \times \Re^n \rightarrow \Re$ eine symmetrische Bilinearform, welche beide durch dieselbe symmetrische Matrix $A = [a_{ik}]$ beschrieben werden. Weiterhin sei $\xi = \sum_i^n x_i \varepsilon_i = (x_1, x_2, \ldots, x_n)$ und es sei $q(\xi)$ beschrieben durch:

$q(\xi) = \dfrac{\phi(\xi,\xi)}{(\xi,\xi)}$, $\xi \neq 0$, wobei

$$\phi(\xi,\xi) = \phi\left(\sum_{i=1}^n x_i \varepsilon_i, \sum_{k=1}^n x_k \varepsilon_k\right) = \sum_{i=1}^n \sum_{k=1}^n x_i x_k \phi(\varepsilon_i, \varepsilon_k) = \sum_{i=1}^n \sum_{k=1}^n x_i x_k a_{ik}$$

16 Satz aus Analysis I: Jede in einem abgeschlossenen beschränkten Intervall stetige Funktion f: $[a, b] \rightarrow \Re$ ist beschränkt und nimmt ihr Maximum und Minimum an.

und $(\xi,\xi) = \sum_{i=1}^{n} x_i^2$. Somit läßt sich q(ξ) schreiben als $q(\xi) = \dfrac{\sum_{i=1}^{n}\sum_{k=1}^{n} x_i x_k a_{ik}}{\sum_{i=1}^{n} x_i^2}$.

Behauptung: Der Vektor ξ ist genau dann ein Eigenvektor von f zum Eigenwert q(ξ), wenn für die partiellen Ableitungen von q gilt:

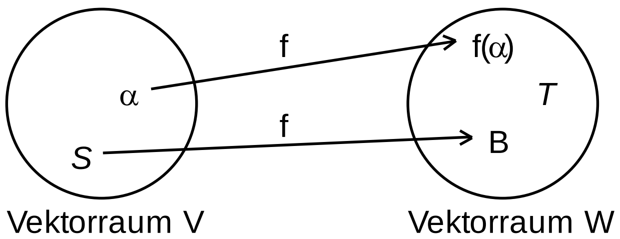
$\dfrac{\partial q(\xi)}{\partial x_j} = 0$ für j = 1, …, n.

4. Folgerungen

4.1 Lineare Abbildungen

Hier soll nun darauf eingegangen werden, was hinter der Berechnung $\underline{y} = (y_1, y_2, \ldots, y_n) = W\underline{x}$ aus Kapitel 2.2 steckt, nachdem ein neuronales System *N* die Eigenvektoren der symmetrischen Korrelationsmatrix C gelernt hat. Es wurde behauptet, daß $(y_1, y_2, \ldots, y_n)$ die Koordinaten des Vektors x bezüglich der orthornormierten Basis $W = (\underline{w}_1, \underline{w}_2, \ldots, \underline{w}_n)$ sind. Wie läuft eine lineare Abbildung nun *technisch* ab?

Sei $f: V \rightarrow W$ eine lineare Abbildung, und es seien $S = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ eine Basis von V und $T = \{\beta_1, \beta_2, \ldots, \beta_m\}$ eine Basis von W. Weiterhin sei $B = \{f(\alpha_1), f(\alpha_2), \ldots, f(\alpha_n)\} = f(S)$.



Lineare Abbildung f

Bezüglich S und T werde f durch die $m \times n$ Matrix A beschrieben. In dieser Matrix stehen spaltenweise die Koordinaten der Bilder der Basisvektoren von V bezüglich der Basis T, also die Koordinaten der Vektoren $B = \{f(\alpha_1), f(\alpha_2), \ldots, f(\alpha_n)\} = f(S)$ bezüglich T.

Soll nun ein Vektor α abgebildet werden, so nimmt man sich diesen Vektor und wendet f darauf an. Man erhält ohne weitere Zwischenschritte $f(\alpha)$. Wie wird eine Abbildung mit Hilfe der Matrix A durchgeführt? Dazu berechnet man die Koordinaten des Vektors α bezüglich der Basis S (durch Lösen eines linearen Gleichungssystems). Nun multipliziert man die Koordinaten $\underline{x} = (x_1, x_2, \ldots, x_n)$ mit der Matrix A und erhält die Koordinaten $\underline{y} = (y_1, y_2, \ldots, y_m)$ des Vektors x bezüglich der Basis T. Jetzt berechnet man die Linearkombination dieser Koordinaten mit den Basisvektoren aus T und erhält so den Vektor $f(\alpha)$.

Wie man leicht sieht, leistet die Matrix A einer Abbildung f nicht das, was die Abbildung selbst leistet, sondern führt nur die Koordinatentransformation $A\underline{x} = \underline{y}$ durch.

Also kann die Berechnung $W\underline{x} = \underline{y}$ aus Kapitel 2.2 streng genommen keine Abbildung sein, da es sich hier bei $\underline{x}$ um einen echten Vektor ($=\alpha$) handelt und *nicht* um Koordinaten eines Vektors bezüglich einer Basis. Zudem wurde behauptet, daß $\underline{y}$ die Koordinaten des Vektors $\underline{x}$ bezüglich der orthonormierten Basis W sind. Bei einer Koordinatentransformation ($A\underline{x} = \underline{y}$ wie oben) handelt es bei der Matrix A aber um eine Transformationsmatrix und nicht gleichzeitig um die Basis, in welcher der Vektor $\underline{x}$ die Koordinaten $\underline{y}$ hat.

Es handelt sich bei W$\underline{x}$ =$\underline{y}$ aber dennoch um eine korrekte Koordinatentransformation. Es besteht lediglich der Unterschied darin, daß man einen *Vektor* vorgibt und *Koordinaten* herausbekommt. Nach Kapitel 3.10 gilt folgende Aussage:

Satz: Es sei V ein endlich-dimensionaler Vektorraum mit Skalarprodukt, und es sei $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ eine orthonormierte Basis. Dann gilt für jedes $\beta \in V$: $\beta = \sum_i^n (\beta, \alpha_i)\, \alpha_i$.

Hat man eine orthonormierte Basis gegeben (wie $W = (\underline{w}_1, \underline{w}_2, \ldots, \underline{w}_n)$), dann kann man die Koordinaten eines Vektors α bezüglich dieser Basis bestimmen, indem man das Skalarprodukt des Vektors α mit den Basisvektoren bildet. Man bekommt also $x_1 = (\alpha, \underline{w}_1)$, $x_2 = (\alpha, \underline{w}_2)$, …, $x_n = (\alpha, \underline{w}_n)$.

Man erhält also die Koordinaten des Vektors $\underline{x}$ aus Kapitel 2.2 durch eine Hintereinanderausführung des Standard-Skalarprodukts durch W$\underline{x}$ =$\underline{y}$. Es handelt sich hier also bei W auf gar keinen Fall um die Beschreibung einer linearen Abbildung im üblichen Sinne. Streng genommen kann jede Funktion $f: V \times V \rightarrow \Re$, die ein Skalarprodukt definiert, dazu verwendet werden, die Koordinaten $\underline{y}$ zu berechnen, so daß man nicht auf die Methode W$\underline{x}$ =$\underline{y}$ beschränkt ist.

4.2 Eigenvektoren

Wie in Kapitel 2.1 beschrieben, existiert ein neuronales System *N*, welches die Eigenvektoren der Korrelationsmatrix C lernt. Die Korrelationsmatrix $[c_{ij}]$ enthält an der Stelle (i, j) die Kovarianz der Eingabekomponenten (x[i], x[j]), wobei der Mittelwert $\langle x[i] \rangle_x$ für alle i gleich Null ist. Ist die Kovarianz zweier Zufallsgrößen x[i], x[j] gleich Null, dann heißen die Größen x[i] und x[j] unkorreliert, andernfalls korreliert.

Wichtig an der Korrelationsmatrix C ist hier aber nur die Tatsache, daß sie symmetrisch ist. Allgemein kann man dann sagen, daß das neuronale System *N* die Eigenvektoren einer symmetrischen Matrix C berechnen kann. Die Eingaben müssen nur so definiert werden, daß sie gerade die Korrelation der Matrix C besitzen[17].

Nach Kapitel 3.13 wird jede selbstadjungierte Selbstabbildung bezüglich einer orthonomierten Basis durch eine symmetrische Matrix C beschrieben. Umgekehrt beschreibt jede symmetrische Matrix C eine selbstadjungierte Selbstabbildung bezüglich einer orthonormierten Basis.

Weiterhin wird in Kapitel 3.13 gezeigt, daß die normierten Eigenvektoren einer selbstadjungierten Selbstabbildung eine orthonormierte Basis eines euklidischen Vektorraums bilden.

Hier soll nun die Behauptung aufgestellt werden, daß die neuronalen Systeme, welche eine vollständige Eigenvektorzerlegung durchführen, genau die Eigenvektoren einer selbstadjungierten Abbildung finden. Diese Eigenvektoren sind orthogonal und normiert und bilden daher einer orthonormierte Basis eines euklidischen Vektorraums.

[17] Es wäre interessant ein Verfahren zu entwickeln, daß zu einer symmetrischen Matrix C eine Menge von Eingabevektoren berechnet, deren Korrelation sich exakt in der Matrix C wiederspiegelt.

Es stellt sich daher die Frage, ob Forschungsergebnisse über selbstadjungierte Abbildungen auf die Eigenvektorzerlegung der neuronalen Netze angewendet werden können. So könnte man die Eigenvektorzerlegung aus einem anderen Blickwinkel betrachten.

Wie in Kapitel 3.16 gezeigt, läßt sich mit Hilfe von selbstadjungierten Selbstabbildungen und quadratischen Bilinearformen recht einfach eine Funktion beschreiben, die die Eigenvektoren einer symmetrischen Matrix C als Extrema besitzt. Es könnte doch nun möglich sein, eine Funktion q: $\mathfrak{R}^n \to \mathfrak{R}$ mit denselben Hilfsmitteln zu konstruieren, welche nur die *normierten* Eigenvektoren einer symmetrischen Matrix als Extrema besitzt.

Wäre das dann nicht die Funktion, auf deren Extrema die Gradientenalgorithmen der neuronalen Systeme *N* zum *Stillstand* kommen?

Dieser Text wurde erstellt mit

Word für Windows 2.0

unter der grafischen Benutzeroberfläche

Michael Rippl, Kalbacher-Straße 5, 6380 Bad Homburg

Windows 3.1

Windows 3.1

Windows 3.1

Michael Rippl, Kalbacher-Straße 5, 6380 Bad Homburg