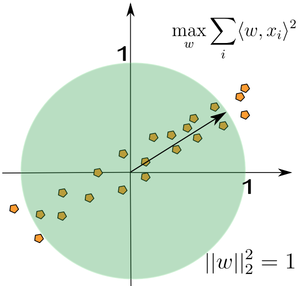
Polynome, Eigenwerte & Eigenvektoren

Irene Winkler & Nico Görnitz

Arbeitsgruppe Maschinelles Lernen

Bsp: Hauptkomponentenanalyse

- Eigenvektoren bilden Lösung eines wichtigen Klasse von Optimierungsproblemen:
$$\max_{\|\mathbf{w}\|=1} \sum_i \langle \mathbf{w}, \mathbf{x}_i \rangle^2 = \sum_i \mathbf{w}^T x_i x_i^T \mathbf{w} = \mathbf{w}^T (\sum_i x_i x_i^T) \mathbf{w} = \mathbf{w}^T C \mathbf{w}$$
- Langrange: $\max_\lambda \min_\mathbf{w} L(\mathbf{w}, \lambda) = -\mathbf{w}^T C \mathbf{w} + \lambda(\|\mathbf{w}\|_2^2 - 1)$
- Ableitung gibt uns das Minimum $0 = \frac{\partial L(\mathbf{w},\lambda)}{\partial \mathbf{w}} = -2C\mathbf{w} + 2\lambda\mathbf{w}$
- daraus folgt: $C\mathbf{w} = \lambda_{max}\mathbf{w}$



Definition & Eigenschaften

- Ein *Polynom* mit Koeffizienten aus einem Körper K hat die Gestalt: $f(x) = \sum_{i=0}^{n} a_i x^i$
- Wenn alle $a_i = 0$ sind, dann spricht man vom *Nullpolynom*
- Der *Grad* von f ist n (d.h. deg$f = n$ und $-\infty$ wenn es das Nullpolynom ist)
- f heisst *normiert*, wenn $a_n = 1$ ist
- sind f,g Polynome, dann ist $\deg(f \cdot g) = \deg f + \deg g$
- Die Anzahl der Nullstellen von f ist kleiner gleich als sein Grad
- Sei λ Nullstelle von f, so gibt es ein eindeutiges g: $f = (x - \lambda) \cdot g$ mit $\deg f = \deg g + 1$

Fundamentalsatz

Der *Fundamentalsatz der Algebra* besagt, dass ein komplexes Polynom vom Grad n größer oder gleich 1 mindestens eine komplexe Nullstelle hat.

Es hat genau n Nullstellen, wenn die Nullstellen entsprechend ihrer Vielfachheit gezählt werden, beispielsweise hat das Polynom $(x-2)^2$ eine doppelte Nullstelle bei $x=2$. Jedes Polynom positiven Grades lässt sich daher in ein Produkt von Linearfaktoren zerlegen.

Definition

Definition. Ein *Eigenvektor* einer quadratischen Matrix $A \in \mathbb{R}^{n \times n}$ zum *Eigenwert* $\lambda \in \mathbb{C}$ ist ein Vektor $\mathbf{v} \in \mathbb{R}^n \setminus \{0\}$ der durch A auf ein Vielfaches $\lambda\mathbf{v}$ von sich selbst abgebildet wird:

$$A\mathbf{v} = \lambda\mathbf{v}$$

Berechnung

Sei $A \in \mathbb{R}^{n \times n}$ eine quadratische Matrix. Dann gilt für jeden Eigenvektor $\mathbf{v}$

$$A\mathbf{v} = \lambda\mathbf{v} \Leftrightarrow (A - \lambda I)\mathbf{v} = 0$$

$\Rightarrow (A - \lambda I)$ muss singulär sein

1. Berechne die Eigenwerte als Nullstellen des *charakteristischen Polynoms* $P(\lambda) := \det(A - \lambda I)$. Dieses Polynom besitzt Grad n.
2. Bestimme für jeden der ermittelten reellen Eigenwerte λ_i eine Basis des Vektorraums $\{\mathbf{v} \in \mathbb{R}^n | (A - \lambda_i I)\mathbf{v} = 0\}$.

Weitere Eigenschaften

Sei $A \in \mathbb{R}^{n \times n}$ eine quadratische Matrix. Es gilt:

- Es gibt höchstens n reelle Eigenwerte und höchstens n linear unabhängige Eigenvektoren.
- Es muss nicht n linear unabhängige Eigenvektoren geben, selbst wenn es n reelle Nullstellen des characteristischen Polynoms gibt.
- Eigenvektoren zu paarweise verschiedenen Eigenwerten sind linear unabhängig.
- A besitzt n paarweise verschiedene Eigenwerte $\Rightarrow$ A besitzt n linear unabhängige Eigenvektoren

Ps: Praktische Berechnung von $A\mathbf{x} = \lambda\mathbf{x}$ durch z.B. QR-Zerlegung oder LU-Zerlegung der Matrix A.

Diagonalisierbarkeit

Definition. Eine quadratische Matrix $A \in \mathbb{R}^{n \times n}$ heißt *diagonalisierbar* falls es eine invertierbare Matrix $S \in \mathbb{R}^{n \times n}$ und eine Diagonalmatrix $\Lambda \in \mathbb{R}^{n \times n}$ gibt, so dass

$$\Lambda = S^{-1} A S.$$

$A = S \Lambda S^{-1}$ nennt man auch *Eigenzerlegung* von A.

Es gilt:

- A besitzt n linear unabhängige Eigenvektoren $\Leftrightarrow$ A diagonalisierbar
- Die Spalten von S sind Eigenvektoren von A, die Diagonale von Λ enthält die zugehörigen Eigenwerte.

Eigenschaften diagonalisierbarer Matrizen

Sei $A \in \mathbb{R}^{n \times n}$ diagonalisierbar mit $A = S\Lambda S^{-1}$.

- Alle Eigenwerte $\neq 0 \Leftrightarrow A$ invertierbar und

$$A^{-1} = (S\Lambda S^{-1})^{-1} = S\Lambda^{-1}S^{-1}$$

- Diagonalisierung erleichtert Potenzierung von A:

$$A^p = S\Lambda^p S^{-1} \text{ für } p \in \mathbb{N}$$

- Die Determinante ist das Produkt der Eigenwerte.
- Die Spur ist die Summe der Eigenwerte.

Spezialfall: Symmetrische Matrizen

Für jede symmetrische Matrix $A \in \mathbb{R}^{n \times n}$ gilt

- Die Eigenwerte sind reell.
- Eigenvektoren zu paarweise verschiedenen Eigenwerten sind orthogonal.
- Es lassen sich immer n orthogonale Eigenvektoren finden.
- A lässt sich zerlegen in

$$A = U \Lambda U^T$$

wobei U eine orthogonale Matrix mit Eigenvektoren als Spalten und Λ die Diagonalmatrix mit den zugehörigen Eigenwerten auf der Diagonalen.

Positiv-Definitheit

Definition. Eine quadratische Matrix $A \in \mathbb{R}^{n \times n}$ heißt

positiv definit	falls $\mathbf{v}^T A \mathbf{v} > 0$
positiv semidefinit	falls $\mathbf{v}^T A \mathbf{v} \geq 0$
negativ definit	falls $\mathbf{v}^T A \mathbf{v} < 0$
negativ semidefinit	falls $\mathbf{v}^T A \mathbf{v} \leq 0$
	für alle $\mathbf{v} \in \mathbb{R}^n \setminus \{0\}$

Satz. Für jede symmetrische Matrix A gilt:

A positiv definit	$\Leftrightarrow$	alle Eigenwerte > 0
A positiv semidefinit	$\Leftrightarrow$	alle Eigenwerte ≥ 0
A negativ definit	$\Leftrightarrow$	alle Eigenwerte < 0
A negativ semidefinit	$\Leftrightarrow$	alle Eigenwerte ≤ 0

- Die wichtigsten Gleichungen für das Rechnen mit Matrizen:
  K. B. Petersen, M. S. Pedersen (2007) *The Matrix Cookbook.*
  `http://www2.imm.dtu.dk/pubdb/views/publication_details.php?id=3274`