Inhaltsverzeichnis

Learning from Data

Mathematische Methoden des digitalen Zeitalters

1 Optimierung

1.1 Grundlagen der Optimierung

Bei der Optimierung betrachtet man vor allem Funktionen, welche Elemente eines n-dimensionalen Vektorraumes auf reelle Zahlen abbilden:

$$f : \mathbb{R}^n \to \mathbb{R}$$

Beispielhafte Anwendungszwecke wären die Minimierung oder Maximierung der Kosten bzw. der Einnahmen eines Unternehmens. Diese würden dem reellen Funktionswert der Funktion entsprechen und ergäben sich aus den Eigenschaften des Unternehmens wie Gehalt, Anzahl der Mitarbeiter, Marketing-Ausgaben usw., welche man zusammenfassend als Vektor eines n-dimensionalen Vektorraumes darstellen könnte.

1.2 Regression

THEOREM 1 *(Methode der kleinsten Quadrate)* Beispiel für die Lösung eines solchen Minimierungsproblems wäre die Methode der kleinsten Quadrate. Gegeben ist hierbei eine lineare Funktion f und Punkte $P_i = (x_i, y_i)$, welche näherungsweise auf der Geraden liegen:

$$\begin{aligned} f(x) &= ax \\ y_i &= f(x_i) + \epsilon_i \end{aligned}$$

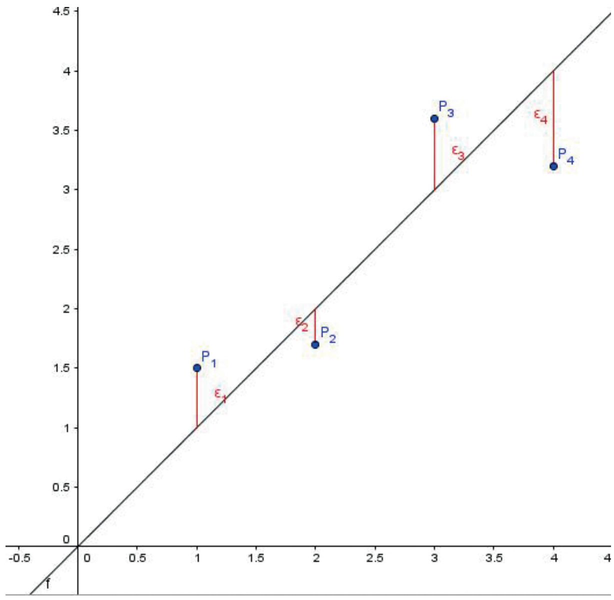

Abb. 1.1: Konvexe Funktion

Wenn nun der Abstand ϵ_i minimiert werden soll, dann gilt:

$$\min h(x) = \frac{\sum_{i=1}^{n}(y_i - ax_i)^2)}{e(y_i, x_i)}$$

Hierbei könnte $e^2, \mid e \mid, h(e)$ exemplarische Möglichkeiten seien. Zur Berechnung dieses Minimums gibt es nun verschiedene Möglichkeiten:

1. Gewöhnliche Tiefpunktberechnung
   z. B. $\quad h'(a) = 0 \quad \wedge \quad h''(a) > 0$
2. Gradientenabstiegsmethode

$a^{t+1} = a^t - \lambda \cdot \delta f(a^t)$
Man findet also ein $\lambda > 0$, sodass $f(a^{t+1}) < f(a^t) \quad | \, \delta f(a^{(t)}) \neq 0$
$g^+(\lambda) = f(a^+ - \lambda \Delta f(a^t))$
$\lambda^+ = argmin(g^+(\lambda))$

DEFINITION 1 *(Lokales und globales Minimum)* Für $f : D \to \mathbb{R}$ ist $x \in D$ ein lokales Minimum, wenn mindestens eine Umgebung N existiert, sodass
$\forall_{y \in N}$ gilt: $f(y) >= f(x)$
$x \in D$ ist dann ein globales Minimum, falls $N = D$.

1.3 Konvexität

DEFINITION 2 *(Konvexe Funktion)* Eine Funktion $f : \mathbb{R}^n \to \mathbb{R}$ ist dann eine konvexe Funktion, falls $\forall_{x,y \in D}$ gilt: $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (a - \lambda) f(y) \quad | \, \forall_{y \in [0,1]}$

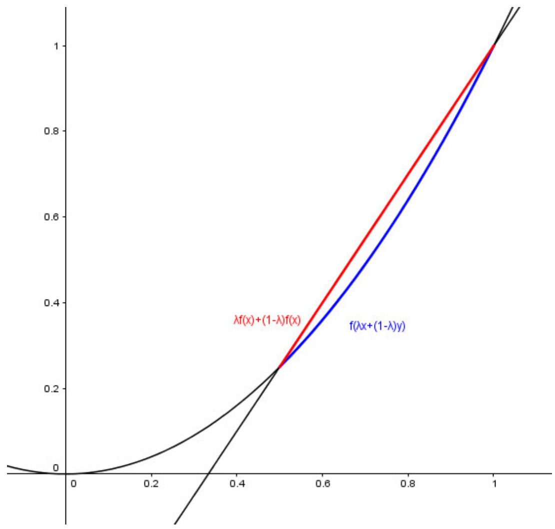

Abb. 1.2: konvexe Funktion

1.3.1 Beispiel zur Konvexität

$$f(u) = \| u \|_2^2 \qquad | \, u \in \mathbb{R}$$
$$f(\lambda u + (1-\lambda)v) = \| \lambda u + (1-\lambda)v \|_2 \leq \| \lambda u \|_2 + \| (1-\lambda)v \|_2 \leq \lambda \| u \|_2 + (1-\lambda) \| v \|_2$$

1.4 Konvexe Mengen

Einführung Im Rahmen der Optimierung von konvexen Funktionen ist es erforderlich, den Begriff der konvexen Menge einzuführen. Um diese Thematik anschaulich darzustellen, verwenden wir zunächst verschiedene geometrische Figuren in Abbildung 1.4, von denen einige konvex sowie andere wiederum nicht konvex sind. Ist es möglich für je zwei beliebige Punkte der Menge eine Verbindungsstrecke zu finden, die selbst ebenfalls in der Menge liegt, so ist die Menge konvex.

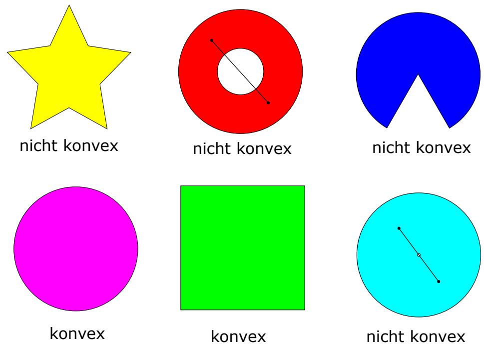

Abb. 1.3: Beispiele für konvexe Mengen

DEFINITION 3 *(Konvexe Menge)*
Eine Menge X heißt konvex, falls für alle $x, y \in X$ und $\lambda \in \mathbb{R}$ und $\lambda \in [0, 1]$ gilt:

$$\lambda x + (1-\lambda) \in X \tag{1.1}$$

Kugel Ein Beispiel für eine konvexe Menge ist die Menge S der Vektoren, die eine Kugel mit dem Radius $r = 1$, deren Mittelpunkt im Ursprung liegt, beschreibt:

$$S = \{\in \mathbb{R}^n | \|v\| \leq 1\}$$

Quadrant Auch die Quadranten des kartesischen Koordinatensystems lassen sich durch eine konvexe Menge erfassen. So gilt z. B. für den ersten Quadranten:

$$L = \{x \in \mathbb{R}^n | x_1 \geq 0, x_2 \geq 0, ..., x_n \geq 0\}$$

Eistüte Bei einer Eistüte bzw. einem quadratischen Kegel handelt es sich um ein weiteres Beispiel für eine konvexe Menge, sonfern für $x \in \mathbb{R}^{n-1}$, $t \in \mathbb{R}$ gilt:

$$Q = \{(x, t) \in \mathbb{R}^n | \|x\|^2 \leq t\}$$

Box Die Menge von Vektoren im $\mathbb{R}$ bilden eine konvexe Menge in Form einer Box mit der Seitenlänge 2, deren Mittelpunkt im Ursprung liegt, wenn gilt:

$$B = \{x \; in \mathbb{R}^n | |x_i| \leq 1, i = 1, ..., n\}$$

1.5 Konvexe Programme

Einführung Bei einem konvexen Programm handelt es sich um eine Funktion f in Abhängigkeit vom Paramenter x. Im Rahmen der mathematischen Optimierung versuchen wir diese Funktion unter Berücksichtigung von Nebenbedingungen zu minimieren. Hierbei sind sowohl die Zielfunktion als auch die Menge der Punkte konvex.

Definition 4 *(Konvexes Programm)* Sei $f : \mathbb{R}^n \to \mathbb{R}$ eine konvexe Funktion und $X \subseteq \mathbb{R}^n$, so minimieren wir die Funktion $f(x)$ unter der Nebenbedigung, dass $x \in X$.

Lineare Programme Lineare Programme lassen sich durch lineare Funktionen beschreiben, z. B. $f(x) = \langle c, x \rangle$. Darüber hinaus sind eine lineare Abbildungsmatrix A, z. B. $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, sowie ein Vektor b und ein weiterer Vektor c, z. B. $c = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, gegeben. x ist dann zulässig, wenn $x \in X$ mit $X : Ax \leq b$ bzw. $X = \{x \in \mathbb{R}^n | Ax \leq b\}$ erfüllt ist. So ergibt beispielsweise die Funktion $f(x)$:

$$f(x) = \langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \rangle = x_1$$

Es folgt für die Nebenbedigungen:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Hieraus folgt: $x_1 \leq 1$, $x_2 \leq 1$.

Quadratische Programme Im Hinblick auf quadratische Programme verwenden wir dieselbe Herangehensweise.
Auch hier sei eine Funktion f gegeben, z. B. $f(x) = x^T Q x$ mit $Q = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, die es zu unter der Bedingung $x \in X$ für $X : Ax \leq b$ zu minmieren gilt.
Hieraus folgt:

$$\begin{aligned} f(x) &= x^T Q x \\ &= (x_1, x_2) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= (x_1, x_2) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= x_1^2 + x_2^2 \end{aligned}$$