

2.4 Verallgemeinerte Ungleichungen

2.4.1 Eigentliche Kegel und verallgemeinerte Ungleichungen

Definition:

Ein Kegel $K \subseteq \mathbf{R}$ heißt eigentlicher Kegel, wenn er die folgenden Bedingungen erfüllt:

- K ist *konvex*

- K ist *abgeschlossen*

- K ist *nichtleer*

- K ist *gerichtet*, das heißt: K darf keine Linien enthalten
gleichbedeutend mit der Aussage:

Definition:

Der eigentliche Kegel wird genutzt um eine *verallgemeinerte Ungleichung* zu definieren, welche eine partielle Ordnung auf \mathbf{R}^n darstellt. Diese besitzt viele Eigenschaften der normalen Ordnung auf \mathbf{R} und wird definiert als:

$$x \preceq_K y \Leftrightarrow y - x \in K$$

Dazu definieren wir eine strikte partielle Ordnung:

$$x \prec_K y \Leftrightarrow$$

wobei aus $x \succ_K y$ folgt $y \preceq_K x$ und aus $x \succ_K y$ folgt $y \prec_K x$.

Für $K=\mathbf{R}_+$ ist die partielle Ordnung \preceq_K die gewöhnliche Ordnung \leq auf \mathbf{R} (Das gleiche gilt für strikt.) Gewöhnliche Ungleichungen sind ein Spezialfall der verallgemeinerten Ungleichungen.

Beispiel 2.14: *Nichtnegativer Orthant und komponentenweise Ungleichheit*
 Der nichtnegative Orthant $K = \mathbf{R}_+^n$ ist ein eigentlicher Kegel. Der verallgemeinerten Ungleichheit mit \preceq_K entspricht die komponentenweise Ungleichheit \leq zwischen Vektoren:

$$x \preceq_K y \text{ heißt, dass } x_i \prec y_i \text{ für } i=1, \dots, n$$

Beispiel 2.15: *Positive semidefinite Kegel und Matrizenungleichungen*

Eigenschaften der verallgemeinerten Ungleichung:

Eine verallgemeinerte Ungleichheit \preceq_K erfüllt viele Eigenschaften wie:

- \preceq_K ist reflexiv
- \preceq_K ist antisymmetrisch
- \preceq_K ist transitiv
- \preceq_K ist verträglich unter Addition
- \preceq_K ist verträglich unter Multiplikation mit einem $\alpha \geq 0$:
- \preceq_K ist verträglich unter der Grenzwertbildung

Die entsprechende strikte verallgemeinerte Ungleichheit erfüllt:

- wenn $x \prec_K y$, dann $x \preceq_K y$
- wenn $x \prec_K y$ und $u \prec_K v$, dann $x+u \prec_K y+v$
- wenn $x \prec_K y$ und $\alpha > 0$, dann: $\alpha x \prec_K \alpha y$
- x kann nicht $\prec_K x$ sein
- wenn $x \prec_K y$, dann gilt für alle hinreichend kleine u und v : $x+u \prec_K y+v$

2.4.2 Minimum und minimale Elemente

Die Schreibweise der verallgemeinerten Ungleichungen soll auf die Analogie zu den gewöhnlichen hindeuten.

Viele Eigenschaften der verallgemeinerten Ungleichungen lassen sich übertragen, einige wichtige jedoch nicht. Zum Beispiel ist \leq eine lineare Ordnung auf \mathbf{R} , das heißt: zwei beliebige Punkte können miteinander verglichen werden, $x \leq y$ oder $y \leq x$. Dies gilt nicht für die verallgemeinerten Ungleichungen.

Dies bereitet Komplikationen, wenn es um Minimum und Maximum geht.

Definition:

Ein $x \in S$ heißt *Minimum* (*Maximum*) der Menge S , wenn für alle $y \in S$ gilt:

$$x \preceq_K y \quad (x \succeq_K y).$$

Wenn eine Menge ein Minimum bzw. Maximum besitzt, so ist dies eindeutig.

Definition:

Ein $x \in S$ heißt ein *minimales* (*maximales*) Element der Menge S , wenn $y \in S$, $y \preceq_K x$ ($y \succeq_K x$) nur für $y = x$ gilt.

Man kann die beiden Zusammenhänge wie folgt beschreiben:

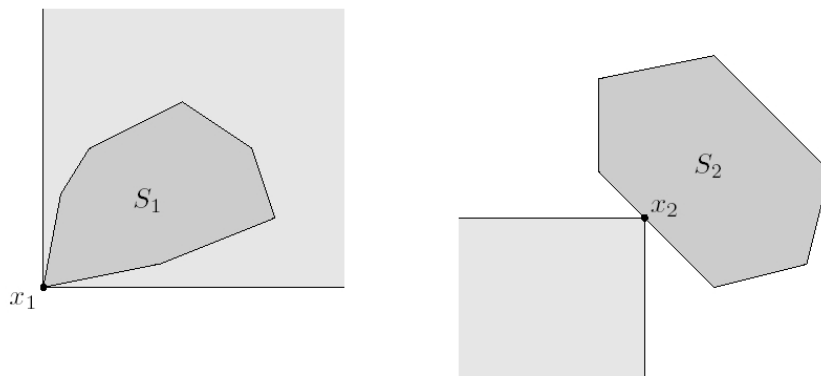
- x ist Minimum von $S \Leftrightarrow$
- x ist ein minimales Element \Leftrightarrow

Beispiel 2.17:

Betrachten den Kegel \mathbf{R}_+^n .

Die Ungleichung $x \leq y$ heißt, dass y sich oberhalb und/oder rechts von x befindet.

- x ist das Minimum von S , wenn alle anderen Punkte aus S oberhalb und rechts von x liegen,
- x ist ein minimales Element von S , wenn kein anderer aus S unterhalb und/oder rechts von x liegt.



2.6 Duale Kegel und verallgem. Ungleichungen

2.6.1 Dualer Kegel

Definition:

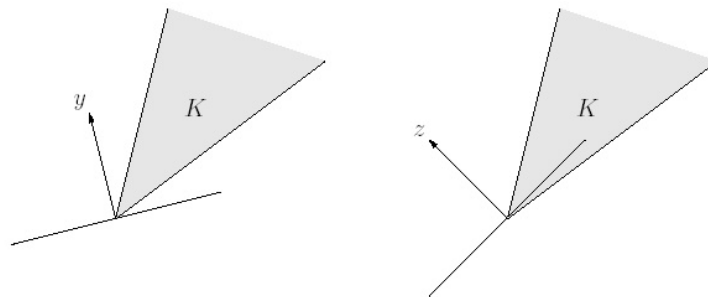
Sei K ein Kegel. Als *dualer Kegel* von K wird die folgende Menge bezeichnet:

$$K^* = \{y: x^T y \geq 0 \text{ für alle } x \in K\}$$

K^* ist ein Kegel und immer konvex, sogar wenn der ursprüngliche Kegel K nicht konvex ist.

Geometrisch:

$y \in K^* \Leftrightarrow -y$ ist der Normalenvektor der Hyperebene, die den Kegel K in seinem Ursprung stützt:



Insbesondere ist K^* der negative Polarkegel K° von K !

Beispiel 2.23 Nichtnegativer Orthant

Der Kegel \mathbf{R}_+^n ist gleichzeitig sein eigener dualer Kegel:

$$y^T x \geq 0 \text{ für alle } x \geq 0 \Leftrightarrow y \geq 0$$

Solche Kegel heißen

Beispiel 2.24: Positiver semidefiniter Kegel

Beispiel 2.25: Duale des Normkegels

Eigenschaften dualer Kegel:

- K^* ist abgeschlossen und konvex
- Aus $K_1 \subseteq K_2$ folgt: $K_2^* \subseteq K_1^*$
- Ist K nichtleer, so ist K^* gerichtet.
- Ist der Abschluss von K gerichtet, so ist K^* nichtleer.
- K^{**} ist der Abschluss der konvexen Hülle von K
daher gilt also, wenn K konvex und abgeschlossen ist, dann

Bemerkung:

Aus diesen Eigenschaften folgt: wenn K ein eigentlicher Kegel ist, so ist K^* auch ein solcher und es gilt: $K^{**} = K$.

2.6.2 Duale verallgemeinerte Ungleichungen

Definition:

Sei K ein eigentlicher konvexer Kegel, also induziert er eine verallgemeinerte Ungleichheit \preceq_K . Dann ist sein dualer Kegel K^* ebenfalls eigentlich und induziert deshalb auch eine verallgemeinerte Ungleichheit.

Wir bezeichnen die verallgemeinerte Ungleichheit \preceq_{K^*} als die zu \preceq_K *duale* verallgemeinerte Ungleichheit.

Wichtige Eigenschaften verallgemeinerter Ungleichheit und ihrer dualen:

- $x \preceq_K y \Leftrightarrow \lambda^T x \leq \lambda^T y$ für alle $\lambda \succeq_{K^*} 0$,
- $x \prec_K y \Leftrightarrow \lambda^T x < \lambda^T y$ für alle $\lambda \succeq_{K^*} 0, \lambda \neq 0$.

Da $K=K^{**}$, ist die zu \preceq_{K^*} duale Ungleichheit gerade \preceq_K , also bleiben die Eigenschaften von einer verallgemeinerten Ungleichheit und ihrer dualen bestehen, falls diese vertauscht werden.

Bemerkung:

Mit diesen Eigenschaften kann man ein verallgemeinertes Farkas Lemma beweisen.

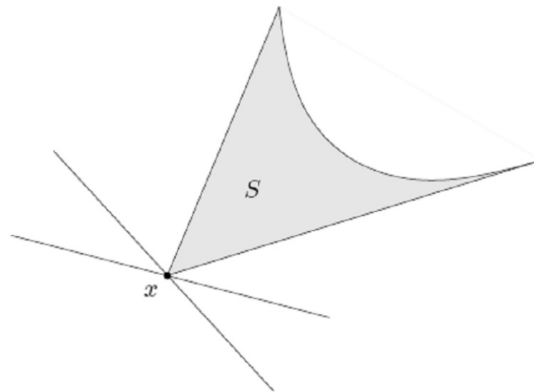
2.6.3 Minimum und minimale Elemente mittels dualen Ungleichungen

Duale Charakterisierung vom Minimum

Definition:

Sei $x \in S$ ein *Minimum* von S , in Bezug auf die verallgemeinerte Ungleichheit $\preceq_K \Leftrightarrow \forall \lambda \succ_{K^*} 0$ x ein *eindeutiger Minimierer* von $\lambda^T z$ über $z \in S$ ist.

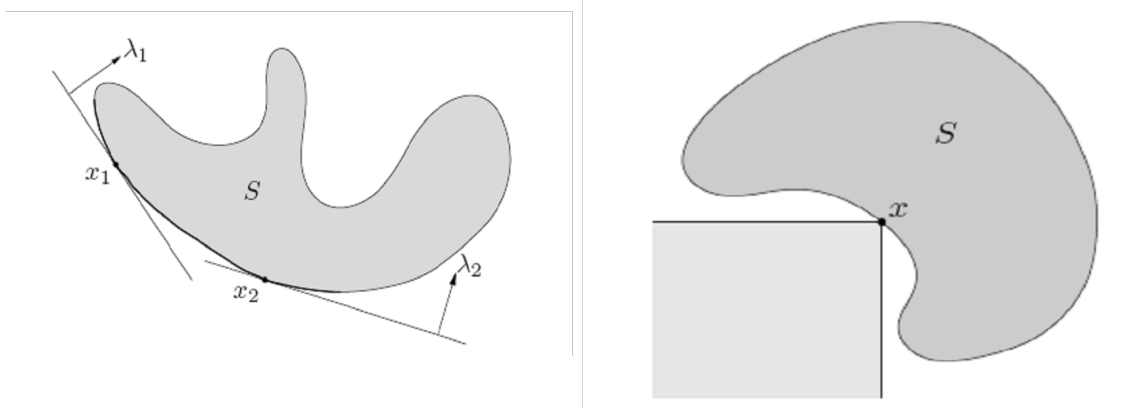
Geometrisch heißt dies, dass für alle $\lambda \succ_{K^*} 0$ die Hyperebene $\{z \mid \lambda^T(z - x) = 0\}$ eine strikt stützende Hyperebene zu S an der Stelle x ist.



Duale Charakterisierung von minimalen Elementen

Hinreichende Bedingung:

Sei $\lambda \succ_{K^*} 0$ und $x \in S$ minimiert $\lambda^T z$ über $z \in S$, dann ist x ein minimales Element. (Die Rückrichtung gilt im Allgemeinen nicht!)

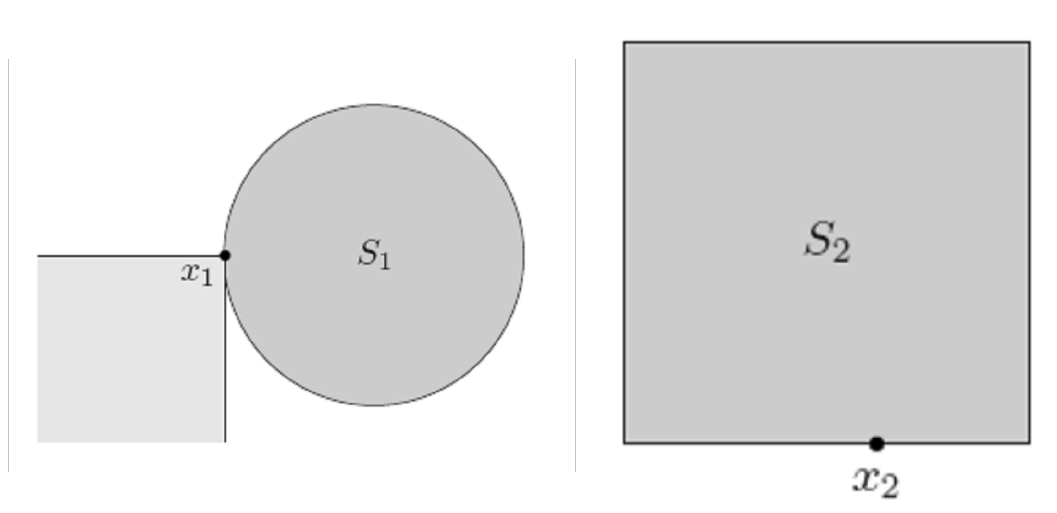


Notwendige Bedingung:

Sei S eine konvexe Menge, dann gilt für jedes minimale Element x , dass ein $\lambda \succeq_{K^*} 0$ (λ ist nicht der Nullvektor) existiert, so dass x der Minimierer von $\lambda^T z$ über $z \in S$ ist.

Beweis:

Sei x minimal $\Leftrightarrow ((x-K) \setminus \{x\}) \cap S = \emptyset$. Unter Anwendung des Theorems der trennenden Hyperebene zwischen konvexen Mengen $(x-K) \setminus \{x\}$ und $S \Rightarrow \exists \lambda \neq 0$ und μ so, dass $\lambda^T(x-y) \leq \mu \forall y \in K$ und $\lambda^T z \geq \mu \forall z \in S$. Aus der 1. Ungleichung folgern wir, dass $\lambda \succeq_{K^*} 0$ ist. Wenn $x \in S$ und $x \in x-K \Rightarrow \lambda^T x = \mu$, die zweite Ungleichung impliziert, dass μ der minimale Wert von $\lambda^T z$ über z ist. $\Rightarrow x$ minimiert $\lambda^T z$ über S , wobei $\lambda \neq 0, \lambda \succeq_{K^*} 0 \Rightarrow$ Behauptung.



Beispiel 2.27: Pareto-optimale Produktionsgrenze

Wir betrachten ein Produkt, welches aus n Ressourcen (z.B.: Arbeit, Elektrizität, Wasser, Material,...) hergestellt wird. Dieses kann auf unterschiedlichen Wegen produziert werden. Jede Produktionsmethode kann durch einen Ressourcenvektor $x \in \mathbf{R}^n$ dargestellt werden, wobei x_i die Menge der Ressource i bezeichnet, die bei dieser Methode verbraucht wird.

Wir nehmen an, dass $x_i \geq 0$ ist und die Inputfaktoren kostbar sind (geringer Verbrauch wird bevorzugt).

Die Produktionsmenge $P \subseteq \mathbf{R}^n$ wird als Menge aller Ressourcenvektoren bezeichnet, die irgendeiner Produktionsmethode entspricht.

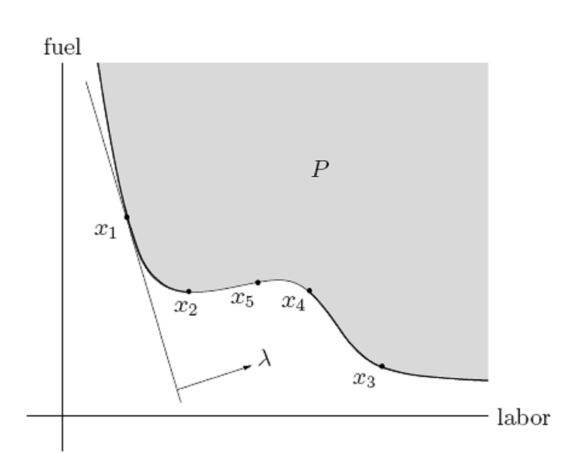
Produktionsmethoden, deren Inputfaktoren minimale Elemente von P sind, werden Pareto-optimal oder effizient genannt. Die Menge aller minimalen Elemente von P heißt effiziente Produktionsgrenze.

Wir sagen, dass eine Produktionsmethode mit dem Vektor x besser als eine andere mit dem Vektor y ist, wenn $x_i \leq y_i \forall i$ und für mindestens ein i $x_i < y_i$ ist. Das entspricht $x \preceq_K y$ für $x \neq y$. Eine Produktionsmethode ist Pareto-optimal bzw. effizient, wenn es keine andere bessere Methode gibt.

Wir finden solche effizienten Produktionsmethoden (minimale Elemente von P) durch:

Minimiere $\lambda^T x = \lambda_1 x_1 + \dots + \lambda_n x_n$ über P , wobei $\lambda > 0$ (λ_i beinhaltet den Preis der Ressource i)

Dies führt zu einer optimalen Produktionsmethode für den Kostenvektor λ . Solange die Preise positiv sind, ist die optimale Produktionsmethode garantiert effizient.



P ist die Produktionsmenge von einem aus Arbeit und Brennstoff hergestelltem Produkt. Die 2 dunklen Kurven veranschaulichen die effiziente Produktionsgrenze. x_1 , x_2 , x_3 sind effizient, x_4 und x_5 sind dies nicht. Der Punkt x_1 ist also die am wenigsten Kosten verursachende Produktionsmethode mit dem Kostenvektor λ . Der Punkt x_2 ist effizient, aber kann nicht durch Minimierung der totalen Kosten $\lambda^T x$ für irgendein $\lambda \geq 0$ gefunden werden.

3.6 Konvexität bezüglich verallgemeinerter Ungleichheit

3.6.1 Monotonie bezüglich einer verallgemeinerten Ungleichheit

Definition:

Sei $K \subseteq \mathbf{R}^n$ ein eigentlicher Kegel mit der damit verbundenen verallgemeinerten Ungleichheit \preceq_K . Eine Funktion $f: \mathbf{R}^n \rightarrow \mathbf{R}$ wird K -nicht fallend genannt, wenn $x \preceq_K y \Rightarrow f(x) \leq f(y)$ und K -steigend, wenn $x \preceq_K y, x \neq y \Rightarrow f(x) < f(y)$

Analog werden K -nicht steigende und K -fallende Funktionen definiert.

Beispiel 3.45: Monotone Vektorfunktionen

Beispiel 3.46: Monotone Matrixfunktionen

Eine Funktion $f: \mathbf{S}^n \rightarrow \mathbf{R}$ wird matrixmonoton (steigend, fallend) genannt, wenn sie monoton bezüglich des positiv semidefiniten Kegels ist.

Einige Beispiele für die Variable $X \in \mathbf{S}^n$:

- $\text{tr}(WX)$, wobei $W \in \mathbf{S}^n$, ist Matrix-nicht fallend, wenn $W \succeq 0$ und Matrix-steigend, wenn $W \succ 0$. Analog für Matrix-nicht steigend und Matrix-fallend.
- $\text{tr}(X)^{-1}$ ist Matrix-fallend auf \mathbf{S}_{++}^n .
- $\det X$ ist Matrix-steigend auf \mathbf{S}_{++}^n und Matrix-nicht fallend auf \mathbf{S}_+^n .

Gradientenbedingungen für Monotonie

Erinnerung:

Satz:

Eine differenzierbare Funktion f mit konvexem Domain ist K -nicht fallend \Leftrightarrow

$$\nabla f(x) \succeq_{K^*} 0 \quad \forall x \in \text{dom } f$$

(Der Gradient muss nichtnegativ in Bezug auf der dualen Ungleichheit sein.)

und sie ist K -steigend, wenn $\nabla f(x) \succ_{K^*} 0 \quad \forall x \in \text{dom } f$ ist. (Die Umkehrung gilt nicht!)

Beweis:

Annahme: Die Funktion f erfüllt $\nabla f(x) \succeq_{K^*} 0 \forall x$ aber f ist nicht K -fallend. $\Rightarrow \exists x, y$ mit $x \preceq_K y \Rightarrow f(y) < f(x)$. Da f differenzierbar, existiert ein $t \in [0,1]$ mit $\frac{d}{dt} f(x+t(y-x)) = \nabla f(x+t(y-x))^T (y-x) < 0$. Wenn $y-x \in K \Rightarrow \nabla f(x+t(y-x)) \notin K^*$ ist. Dies ist jedoch ein Widerspruch zu der Annahme, dass $\nabla f(x) \succeq_{K^*} 0$ überall gilt. Analog wird K -steigend bewiesen.

3.6.2 Konvexität bezüglich einer verallgemeinerten Ungleichheit

Definition:

Sei $K \subseteq \mathbf{R}^n$ ein eigentlicher Kegel mit der damit verbundenen verallgemeinerten Ungleichheit \preceq_K . Wir sagen, $f: \mathbf{R}^n \rightarrow \mathbf{R}^m$ ist K -konvex $\forall x, y$ und $0 \leq \theta \leq 1$, wenn gilt:

$$f(\theta x + (1-\theta)y) \preceq_K \theta f(x) + (1-\theta)f(y)$$

und *strikt* K -konvex, wenn gilt:

$$f(\theta x + (1-\theta)y) \prec_K \theta f(x) + (1-\theta)f(y) \forall x \neq y \text{ und } 0 < \theta < 1$$

Diese Definitionen reduzieren sich zur normalen und strikten Konvexität, wenn $m = 1$ und $K = \mathbf{R}_+$

Beispiel 3.47: Konvexität bezüglich komponentenweiser Ungleichheit

Eine Funktion $f: \mathbf{R}^n \rightarrow \mathbf{R}^m$ ist konvex bzgl. komponentenweiser Ungleichheit $\Leftrightarrow \forall x, y$ und $0 \leq \theta \leq 1: f(\theta x + (1-\theta)y) \preceq_K \theta f(x) + (1-\theta)f(y)$ für jede Komponente f_i gilt (d.h. jede Komponente f_i ist konvex).

Die Funktion f ist strikt konvex \Leftrightarrow

Beispiel 3.48: Matrixkonvexität

Sei f eine symmetrische matrixwertige Funktion, z.B.: $f: \mathbf{R}^n \rightarrow \mathbf{S}^m$. Die Funktion f ist konvex bzgl. der Matrixungleichheit, wenn

$$f(\theta x + (1-\theta)y) \preceq \theta f(x) + (1-\theta)f(y) \forall x, y \text{ und } \theta \in [0,1]$$

wird häufig Matrixkonvexität genannt. Eine äquivalente Definition ist, dass die skalare Funktion $z^T f(x) z$ für alle Vektoren z konvex ist. Analog wird strikte Matrixkonvexität definiert.

Einige Beispiele:

- Die Funktion $f(X) = XX^T$ mit $X \in \mathbf{R}^{n \times m}$ ist matrixkonvex, wenn für fixierte z die Funktion $z^T XX^T z = \|X^T z\|_2^2$ eine konvexe quadratische Funktion von X ist.
- Die Funktion X^p ist matrixkonvex auf \mathbf{S}_{++}^n für $1 \leq p \leq 2$ oder $-1 \leq p \leq 0$ und matrixkonkav für $0 \leq p \leq 1$
- Die Funktion $f(X) = e^{-X}$ ist nicht matrixkonvex auf \mathbf{S}^n für $n \geq 2$.

Duale Charakterisierung von K - Konvexität

Eine Funktion f ist K - konvex $\Leftrightarrow \forall \omega \succeq_{K^*} 0$ die reellwertige Funktion $\omega^T f$ konvex ist.
 f ist strikt K - konvex $\Leftrightarrow \forall \omega \succeq_{K^*} 0; \omega \neq 0$ die Funktion $\omega^T f$ strikt konvex ist.
(Diese Aussagen folgen direkt aus der Definition und Eigenschaften der dualen Ungleichheit.)

Differenzierbare K - konvexe Funktionen

Eine differenzierbare Funktion f ist K - konvex genau dann, wenn ihr Domain konvex ist und $\forall x, y \in \text{dom } f$ gilt:

$$f(y) \succeq_K f(x) + Df(x)(y-x)$$

f ist strikt K - konvex $\Leftrightarrow \forall x, y \in \text{dom } f$ und $x \neq y$ gilt:

$$f(y) \succ_K f(x) + Df(x)(y-x)$$

Kompositionstheorem

Viele Ergebnisse von Kompositionen können für die K - Konvexität verallgemeinert werden.

Zum Beispiel : Sei $g : \mathbf{R}^n \rightarrow \mathbf{R}^p$ K - konvex und $h : \mathbf{R}^p \rightarrow \mathbf{R}$ ebenfalls konvex. Wenn die Erweiterung \tilde{h} von h K - nicht fallend ist, so ist $h \circ g$ konvex.

Dies verallgemeinert, wenn man eine nicht fallende konvexe Funktion auf eine konvexe Funktion anwendet, dass die resultierende Funktion eine konvexe Funktion ist.