TECHNISCHE UNIVERSITÄT CHEMNITZ
Fakultät für Wirtschaftswissenschaften
Professur für Volkswirtschaftslehre IV
- Finanzwissenschaft -
Professor Dr. Thomas Kuhn

TECHNISCHE UNIVERSITÄT CHEMNITZ

Skript zur Vorlesung

Mathematische Methoden der Ökonomik

Prof. Dr. Thomas Kuhn

WS 2020/2021

05.10.2020

Inhaltsverzeichnis:

1 Ökonomische Modelle 1
- 1.1 Definitionen 1
- 1.2 Allgemeines Marktgleichgewicht 2

2 Einführung in die Integral- und Differentialrechnung 8
- 2.1 Die Grundidee hinter Infinitesimalen 8
- 2.2 Rechenregeln für Differentialquotienten 10

3 Optimierung ohne Nebenbedingungen 11
- 3.1 Das totale Differential 11
- 3.2 Optimierung 13
- 3.3 Bedingungen zweiter Ordnung 13
  - 3.3.1 Funktionen einer Variable 14
  - 3.3.2 Funktionen zweier Variablen 14
  - 3.3.3 Allgemeine Funktionen 15
  - 3.3.4 Benötigte Konzepte aus der linearen Algebra: 15
  - 3.3.5 Optimierung mit Bedingungen zweiter Ordnung: 16
  - 3.3.6 Vorgehen: 16

4 Optimierung mit Gleichheitsnebenbedingung 18
- 4.1 Satz 1 (Bedeutung des Gradienten): 21
- 4.2 Satz 2: („Darstellungssatz“) 23
- 4.3 Satz 3: (Zuwachs im Gradientenhalbraum) 23
- 4.4 Satz 4: (Orthogonalität des Gradienten der Zielfunktion auf dem zulässigen Bereich im Optimum) 24
- 4.5 Satz 5: (Gradientekombinationen des zulässigen Bereiches Spannen den zum zulässigen Bereich orthogonalen Raum auf) 25
- 4.6 Satz 6: (Lagrange-Verfahren) 27
- 4.7 Zusammenfassung: (alles lokal) 29
- 4.8 Bedingungen zweiter Ordnung 30

5 Karush- Kuhn- Tucker- Verfahren (KKT) 32

6 Homogene Differentialgleichungen 39
- 6.1 Harrod-Domar-Modell ....
- 6.2 Lösen von homogenen Differentialgleichungen

7 Inhomogene Differentialgleichungen 44

7.1 Solow-Wachstums-Modell

7.2 Lösen von inhomogenen Differentialgleichungen

1. Ökonomische Modelle (Prof. Dr. Thomas Kuhn)

Definitionen

Ökonomisches Modell:

besteht aus einem System von Annahmen
Formuliert in der Sprache der Mathematik beinhalten die *Annahmen* in der Regel:

– mathematische Funktionen

– Identitäten, Gleichungen

– Verhaltensannahmen

Mathematische Funktionen bilden i.d.R. Hypothesen über Kausalabhängigkeiten zwischen ökonomischen Variablen ab, z.B. die Abhängigkeit zwischen Preis und nachgefragter Menge eines Gute

Identitäten beschreiben z.B. Gleichgewichtsbedingungen auf Märkten oder Zusammenhänge in der Volkswirtschaftlichen Gesamtrechnung ex-post

Verhaltensannahmen formulieren Zielfunktionen und reflektieren die Präferenz oder Ziele von Wirtschaftssubjekten und Institutionen

Endogene vs. Exogene Variablen:

Endogene Variablen in einem ökonomischen Modell sind die ökonomischen Variablen, die im Modell als Lösungen eines Annahmen-Systems, etwa mathematischen Gleichungssystems oder Optimierungsproblems, bestimmt werden

Exogene Variablen oder Parameter sind ökonomische Größen, deren Werte gegeben sind und nicht im System bestimmt werden, in der Regel können sie mehrere Werte annehmen

im Allgemeinen hängen die Lösungen eines ökonomischen Modells von allen Parametern des Modells ab (es sei denn, eine Variable hebt sich im Verlauf des Lösungsweges auf und ist unter Umständen nicht mehr sichtbar in den Lösungen

Komparative Statik:

Ist ein Verfahren, das die Abhängigkeit der Lösungen von den Parametern untersucht, durch Variation der Parameterwerte Abb. 1.3 (Blick auf Gesamtmarkt)

Allgemeines Marktgleichgewicht

Beispiel für ein ökonomisches Modell

Im einfachsten Fall besteht ein *Modell des Allgemeinen Marktgleichgewichts* aus den folgenden Elementen:

- zwei finale Güter
- ein repräsentativer Konsument
- ein repräsentatives Unternehmen
- zwei Gütermärkte, auf denen die beiden Güter gehandelt werden
- eine begrenzte Ressource mit zwei Verwendungen

Annahmen:

- Konsumenten maximieren Nutzen unter einer Budgetrestriktion
- Anbieter maximieren ihren Gewinn unter der Ressourcenbeschränkung
- Märkte sollen sich simultan im Gleichgewicht befinden, d.h. auf beiden Märkten herrscht ein Gleichgewicht zwischen Güternachfrage und Güterangebot, alle Märkte sind geräumt
- Sozialer Planer maximiert gesellschaftliche Wohlfahrt (Benchmark)

Formulierung des Modells

Mengen der beiden Güter bezeichnet als: x_1, x_2

Repräsentativer Konsument

ist Mengenanpasser: Preise p_1, p_2 und Einkommen y gegeben, besitzt neoklassische Nutzenfunktion $u(x_1, x_2)$,

Konsument maximiert seinen Nutzen, abhängig vom Güterkonsum, unter der Einkommensbeschränkung:

Optimierungsproblem:

$$\begin{cases} u(x_1, x_2) \to \max\limits_{x_1, x_2} \\ s.t.\ p_1 x_1 + p_2 x_2 \le y \end{cases}$$

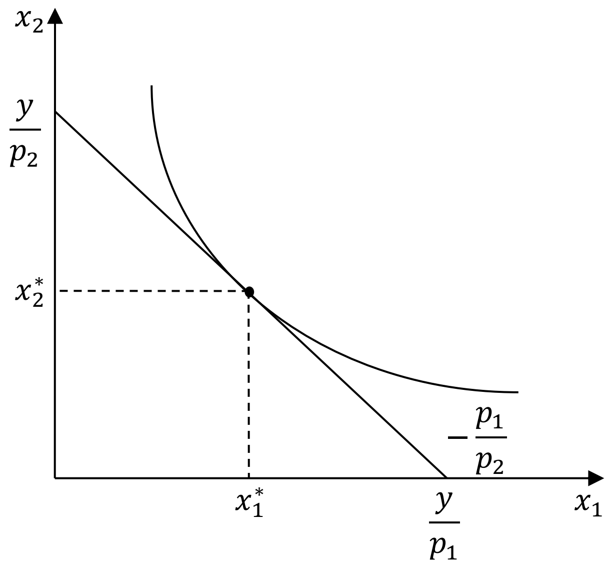
Im Optimum: (x_1^*, x_2^*) gilt bei neoklassischen Nutzenfunktionen:

$$p_1 x_1 + p_2 x_2 = y$$

d. h. das gesamte Einkommen wird aufgebraucht;

Graphische Lösung:

Optimale Güternachfrage befindet sich im Tangentialpunkt von Budgetgeraden und Indifferenzkurve der Nutzenfunktion, dort ist die Marginale Rate der Substitution zwischen Gut 1 und Gut 2, $dx_2(x_1; u^*)/dx_1$, gleich dem Preisverhältnis, wobei $x_2(x_1; u^*)$ eine Auflösung von $u(x_1, x_2) = u^*$ nach x_2 ist (und existiert), wobei u^* das optimale Nutzenniveau bezeichnet.



Lösungen stellen die Güternachfragen allgemein in Abhängigkeit der Preise und des Einkommens dar:

$$x_1^N = x_1^N(p_1, p_2, y)$$

$$x_2^N = x_2^N(p_1, p_2, y)$$

Repräsentatives Unternehmen

- Ist Mengenanpasser,
- Marktpreise gegeben, Ressourcen beschränkt auf R, verwendbar für beide Güter
- Neoklassische Produktionsmöglichkeiten (Opportunitätskosten!)
- Produktionsmöglichkeiten $F(x_1, x_2; R) = 0$ gegeben (PPF) und beschränkt, wegen der Beschränkung der Ressourcen

Kurve der Produktionsmöglichkeiten ist Lösung des Gleichungssystems:

$$\begin{cases} x_1 = x_1(R_1) \\ x_2 = x_2(R_2) \\ R_1 + R_2 = R \end{cases}$$

mit den Produktionsfunktionen $x_1 = x_1(R_1)$ und $x_2 = x_2(R_2)$
Lösung: es gilt:

$x_1 = x_1(R - R_2)$,

$x_1 = x_1\left(R - x_2^{-1}(x_2)\right)$, mit $R_2 = x_2^{-1}(x_2)$ und x_2^{-1} Umkehrfunktion von x_2

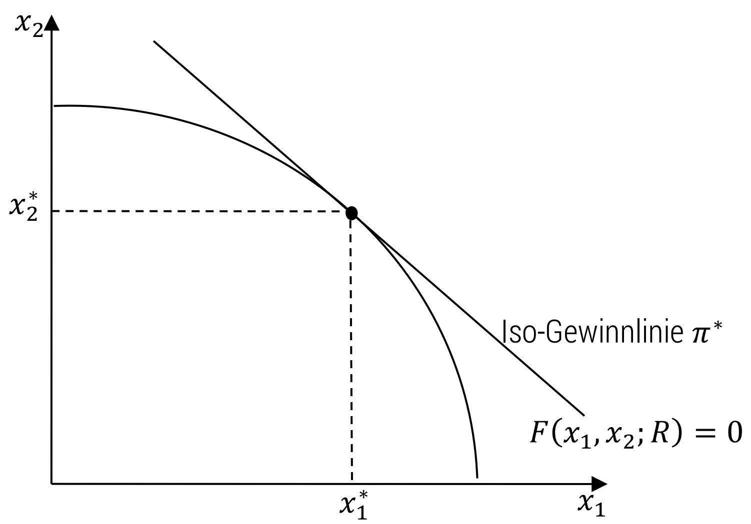
$x_1 = f(x_2; R)$ bzw. $x_1 - f(x_2; R) := F(x_1, x_2; R) = 0$

Optimierungsproblem: beschränkte Maximierung des Gewinns π:

$$\begin{cases} \max\limits_{x_1, x_2} \pi = p_1 x_1 + p_2 x_2 \\ \text{s.t. } F(x_1, x_2; R) = 0 \end{cases}$$

Graphische Lösung:

Optimales Güterangebot befindet sich im Tangentialpunkt von Iso-Gewinnlinie und Kurve der Produktionsmöglichkeiten, dort ist die marginale Rate der Transformation zwischen Gut 1 und 2, $dx_2(x_1; R)/dx_1$, gleich dem Preisverhältnis, wobei $x_2(x_1; R)$ die Auflösung von $F(x_1, x_2; R) = 0$ nach x_2 ist (und existiert)



Angebotsfunktionen allgemein abhängig von Preisen und Ressourcenbeschränkung:

$x_1^A = x_1^A(p_1, p_2; R)$

$x_2^A = x_2^A(p_1, p_2; R)$

Allgemeines Marktgleichgewicht

genügt simultan den Bedingungen

$x_1^N = x_1^A$ (Markt 1)

$x_2^N = x_2^A$ (Markt 2)

$y = \pi$ (Einkommensgleichung)

ausführlich:

$x_1^N(p_1; p_2, y) = x_1^A(p_1; p_2, R)$

$x_2^N(p_2; p_1, y) = x_2^A(p_2; p_1, R)$

$y = \pi(p_1; p_2, R)$

Die *Lösung des Gleichungssystems* ist unter den Annahmen der neoklassischen Ökonomik bezüglich der Produktions- und Nutzenfunktionen eindeutig nur für relative Preise bestimmt, nicht für absolute Preise, d.h. eine Veränderung aller Preise um einen konstanten Faktor $\mu > 0$ ändert das allgemeine Marktgleichgewicht nicht.

Walras Law:

Da die Ressourcen den Unternehmen kostenlos zur Verfügung stehen und es keine Lagerbestände gibt, gilt die Einkommensgleichung:

$y = \pi$ (Einkommen der Konsumenten entspricht der Wertschöpfung (hier Gewinn!) der Unternehmen)

und somit:

$p_1 x_1^N + p_2 x_2^N = p_1 x_1^A + p_2 x_2^A \Leftrightarrow \quad p_1(x_1^N - x_1^A) + p_2(x_2^N - x_2^A) = 0$

$\Leftrightarrow \sum_{i=1}^{n} p_i(x_i^N - x_i^A) = 0$

Walras-Law: Wert aller Überschussnachfragen ist Null, wobei Überschussnachfrage auf Markt *i*: $EX_i = x_i^N - x_i^A$

Falls aber $p_1(p_2, y)$ so bestimmt wird, dass $x_1^N = x_1^A$ gilt (Gleichgewicht), dann muss auch $x_2^N = x_2^A$ für beliebige p_2 gelten, also insbesondere für den Preis p_2, der im Gleichgewicht auf Markt 1 unterstellt ist.

In Partialmodellen kann daher Gut 2 als Numeraire mit $p_2 := 1$ aufgefasst werden und der Wert der Güter in Einheiten von Gut 2 ausgedrückt werden, d.h. für die Relativpreise gilt dann $(p_1, 1)$

Bestimmung der Relativpreise

Das Marktgleichgewicht ist charakterisiert durch:

$$x_1^N = x_1^A$$

$$x_2^N = x_2^A$$

$$\sum_{i=1}^{2} p_i\left(x_i^N - x_i^A\right) = 0 \qquad \text{(Walras)}$$

Im Grunde lautet das Gleichgewichtssystem unter Berücksichtigung der Nachfrage- und Angebotsfunktionen ausführlich:

$$x_1^N(p_1, p_2, y) - x_1^A(p_1, p_2) = 0$$

$$x_2^N(p_1, p_2, y) - x_2^A(p_1, p_2) = 0$$

$\sum_{i=1}^{2} p_i \left(x_i^N(p_1, p_2, y) - x_i^A(p_1, p_2, R)\right) = 0$, (Walras)

Die Nachfragefunktionen wiederum sind Lösungen der FOCs:

$$\frac{\partial u_1}{\partial x_1}(x_1, x_2) - p_1\lambda = 0$$

$$\frac{\partial u_1}{\partial x_2}(x_1, x_2) - p_2\lambda = 0$$

$$p_1x_1 + p_2x_2 = y$$

bzw.

$$\frac{\left(\frac{\partial u_1}{\partial x_1}\right)}{\left(\frac{\partial u_1}{\partial x_2}\right)} = \frac{p_1}{p_2}$$

$$p_1x_1 + p_2x_2 = y$$

Man sieht: Die Nachfragefunktionen sind homogen vom Grad 0 in p_1, p_2 und y, das heißt es gilt für alle $\mu > 0$:

$$x_1^N(p_1, p_2, y) = x_1^N(\mu p_1, \mu p_2, \mu y)$$

und analog

$$x_2^N(p_1, p_2, y) = x_2^N(\mu p_1, \mu p_2, \mu y)$$

Die Angebotsfunktionen sind Lösungen von:

$$p_1 + \lambda \frac{\partial F}{\partial x_1} = 0$$

$$p_2 + \lambda \frac{\partial F}{\partial x_2} = 0$$

$$F(x_1, x_2; R) = 0$$

und ebenfalls homogen vom Grad 0 in den Preisen p_1, p_2, das heißt für alle $\mu > 0$ gilt

$$x_1^A(p_1, p_2) = x_1^A(\mu p_1, \mu p_2)$$

und analog

$$x_2^A(p_1, p_2) = x_2^A(\mu p_1, \mu p_2),$$

denn die Lösung von

$$\frac{\frac{\partial F}{\partial x_1}}{\frac{\partial F}{\partial x_2}} = \frac{p_1}{p_2}$$

$$F(x_1, x_2) = 0$$

ändert sich nicht bei einer Änderung der Preise mit dem Faktor μ.

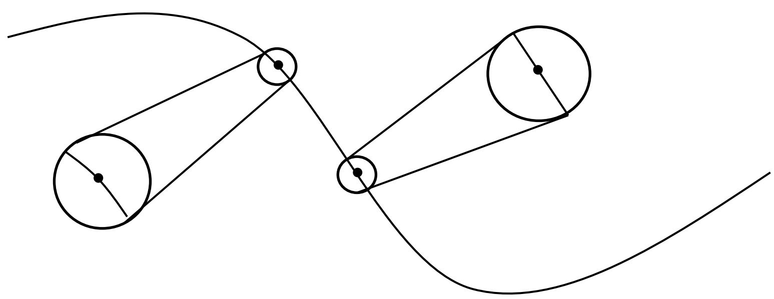
Der Gewinn des Unternehmens π steigt mit dem Faktor μ, denn $\pi(\mu p_1, \mu p_2) = \mu\pi$ und damit steigt auch das Einkommen der Konsumenten mit dem Faktor μ, wie oben bei der Homogenität vorausgesetzt.

In Partialmodellen betrachtet man einen singulären Markt mit Preisverhältnis $\frac{p_1}{p_2}$, bei dem $p_2 = 1$ gesetzt wird und Gut 2 als Numeraire dient.

Komparative Statik in diesem Modellrahmen könnte man durchführen für die exogenen Parameter *R* und *y*

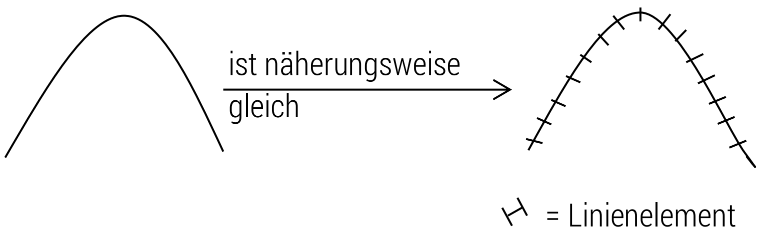
Einführung in die Integral- und Differentialrechnung (Radomir Pestow)

Die Grundidee hinter Infinitesimalen

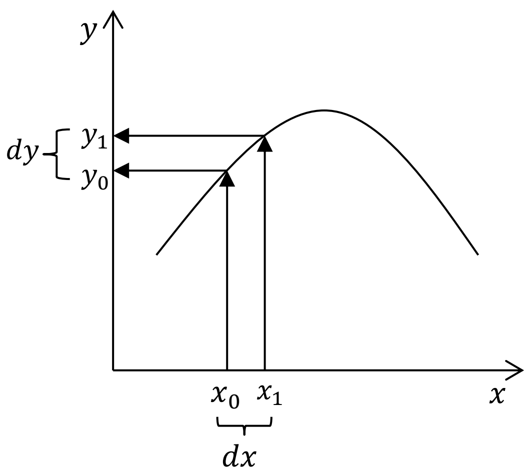


Ein hinreichend kleines Stückchen einer Linie ist annähernd gerade. Ein solches Stück wird auch als Linienelement bezeichnet. Hinreichend klein heißt, dass der Unterschied zwischen Linienstück und echter Gerade für praktische Zwecke vernachlässigt werden kann.

- ➔ Jede Linie im Anschauungsraum lässt sich aus Linienelementen zusammengesetzt denken:



Bedeutung der Infinitesimalberechnung rührt daher, dass man zum Verständnis komplexer Formen lediglich die Elementarformen verstehen muss. Versteht man Strecken, versteht man auch beliebige Linien.
Führen nun Koordinatensystem zum Rechnen ein.



Eine Linie ist eine Funktion, wenn jede Ordinate (Parallelen zur y-Achse) nur in einem Punkt geschnitten wird.

Allgemein: Eine Funktion ist eine Vorschrift die eine abhängige Variable mit einer unabhängigen Variable verknüpft.

$$y \quad = f(x)$$

abh. Variable ↑ unabh. Variable

Funktion

Die Gerade ist charakterisiert durch die Steigung: $\frac{dy}{dx}$

und dem Stützpunkt: x_0

Anmerkung: Tangente fällt in einem hinreichend kleinen Bereich mit dem Linienelement annähernd zusammen.

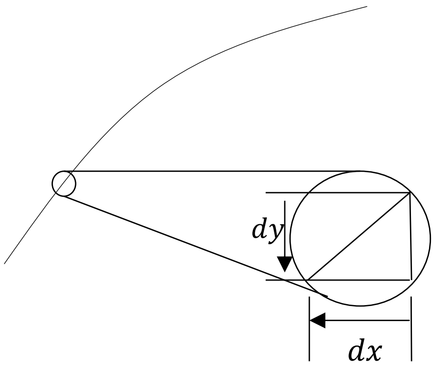
Formal ist die Ableitung definiert durch:

$$\frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = f'(x)$$

Intuitiv kann man Δx als hinreichend klein interpretieren.

Warum läuft beim Grenzübergang nichts schief?

$\frac{dy}{dx} \to \infty$? wegen $\Delta x \to 0$?



Steigerung ändert sich nicht

Wie in der Graphik zu sehen, ändert sich die Steigung im hinreichend kleinen Bereich nicht mehr.

Lokal lässt sich die Änderung des Funktionswertes also so darstellen:

$$dy = f'(x)dx$$

Rechenregeln für Differentialquotienten

Differentialrechnung bestehen im Wesentlichen in der Anwendung folgender Rechenregeln.

Die Funktionen, die uns in dieser Vorlesung begegnen werden, sind algebraische Kombinationen $(+, -, \cdot, \div)$ aus 3 Elementarfunktionen.

Elementarfunktionen und ihre Ableitungen

Potenzfunktion: $ax^m \to amx^{m-1}$

Exponentialfunktion: $ae^{(bx)} \to abe^{(bx)}$

Logarithmus: $a\ln(bx) \to \frac{a}{x}$

Algebraische Kombinationen von Funktionen:

Summe:

$$f(x) + y(x) \to f'(x) + y'(x)$$

Produkt:

$$f(x) \cdot y(x) \to f'(x) \cdot y(x) + f(x) \cdot y'(x)$$

bzw.:

$$f_1(x) \cdot f_2(x) \cdot f_3(x) \to f_1{}'(x) \cdot f_2(x) \cdot f_3(x) + f_1(x) \cdot f_2{}'(x) \cdot f_3(x) + f_1(x) \cdot f_2(x) \cdot f_3{}'(x)$$

Produkt mit Konstante:

$$a \cdot f(x) \to a \cdot f'(x)$$

Funktionsverkettung (Kettenregel):

$$f\big(y(x)\big) \to f'\big(y(x)\big) \cdot y'(x)$$

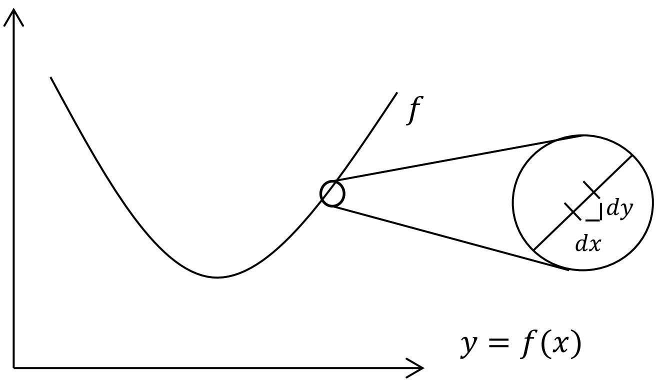
äußere Funktion innere Funktion

Optimierung ohne Nebenbedingungen

Das totale Differential

Leitende Grundidee ist hierbei die (lokale) Linearisierung von Funktionen.

Kennen bereits Linearisierung von Linien, insbesondere Funktionsgraphen



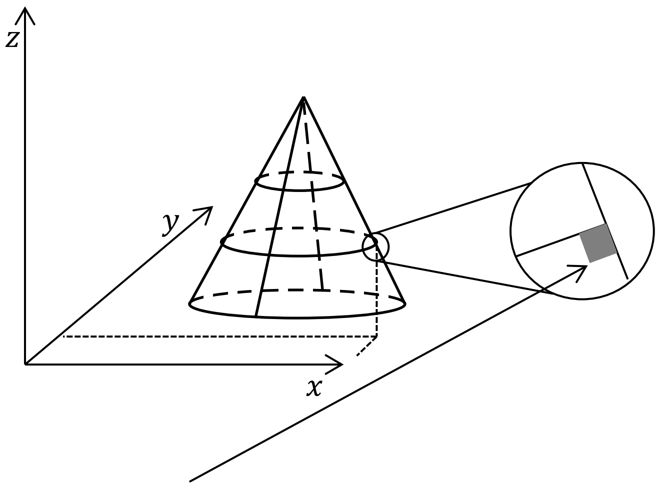
Können Funktionsänderungen lokal durch Differentiale darstellen (einfache lineare Funktionen):

$dy = s \cdot dx$

Stützpunkt

Steigung = $f'(x)$

Wir Verallgemeinern dieser Idee zunächst auf Flächen.



Das Flächenstück ist annähernd eben, falls klein genug. Wir sprechen dann von einem Flächenelement.

- ➔ Flächenelemente lassen sich in der Näherung formelhaft also genauso darstellen, wie Ebenen.

Betrachten Formel für Tangentenebene (Ebene in der unser Flächenelement liegt).

$z = a_1 x + a_2 y + c$

Uns interessiert die Änderung des Funktionswertes in der Nähe zum Stützpunkt, $f(x_0, y_0)$.

Wie ändert sich also f wenn man von (x_0, y_0) zu (x_1, y_1) geht?

$$\begin{aligned} d_z &= z_1 - z_2 \\ &= a_1 x_1 + a_2 y_1 + c - (a_1 x_0 + a_2 y_1 + c) \\ &= a_1(x_1 - x_0) + a_2(y_1 - y_0) \\ &= a_1 dx + a_2 dy \end{aligned}$$

Bestimmen nun a_1 und a_2.

Setzen $dy = 0 \Longrightarrow dz = a_1 dx \Leftrightarrow \frac{dz}{dx} = a_1$

dz ist hierbei Änderung in x- Richtung

das heißt, a_1 ist die Ableitung von f in x Richtung im Punkt (x_0, y_0).

Um Verwechslungen zu vermeiden schreiben wir: $\frac{\partial f}{\partial x}$ oder auch einfach f_x, f_1. Die Zahl im Subindex deutet an, dass hierbei nach dem ersten Argument abgeleitet wird.

Analog für $a_2 = \frac{\partial z}{\partial y} = \frac{\partial f}{\partial y}$

Insgesamt also: $dz = df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dx$

Totales Differential von f in x_0, y_0 (=lineare Näherung für eine Funktionsänderung).

Ganz allgemein:

Eine Funktion beliebig vieler Variablen heißt total differenzierbar, wenn ihre Änderung überall linear angenähert werden kann.

Das heißt, überall gilt:

$$dz = \frac{\partial z}{\partial x_1} dx_1 + \frac{\partial z}{\partial x_2} dx_2 + \cdots + \frac{\partial z}{\partial x_m} dx_m$$

Anders gesagt: die gesamte Änderung dz ist näherungsweise die Summe der Änderung in die jeweiligen Richtungen $\frac{\partial z}{\partial x_i} dx_i$

Wir fassen die partiellen Ableitungen zu einem Vektor (= Reihe von Zahlen) zusammen: dem Gradientenvektor, oder auch einfach nur Gradient.

$$\nabla \mathrm{z} = \begin{pmatrix} \frac{\partial z}{\partial x_1} \\ \vdots \\ \frac{\partial z}{\partial x_m} \end{pmatrix}$$

Optimierung

$\max f(x)$
$x \in \mathbb{R}^m$

Wollen also f maximal, d.h. größtmöglich machen.

Im einem Maximum (x_0, y_0) muss gelten:

$\frac{\partial f}{\partial x_i} = 0$, denn sonst wäre $df = \frac{\partial f}{\partial x_i} dx_i$ mit $dx_j = 0 \qquad j \neq i$, für irgendein i.

das heißt, f ließe sich in (x_0, y_0) weiter vergrößern (mit $dx_i > 0$ falls $\frac{\partial f}{\partial x_i} > 0$ und umgekehrt). Analoges gilt aber auch im Minimum und auch in Sattelflächen zum Beispiel. Das Kriterium ist also lediglich notwendig, aber nicht hinreichend.

Bedingungen zweiter Ordnung

Wir haben gesehen, dass

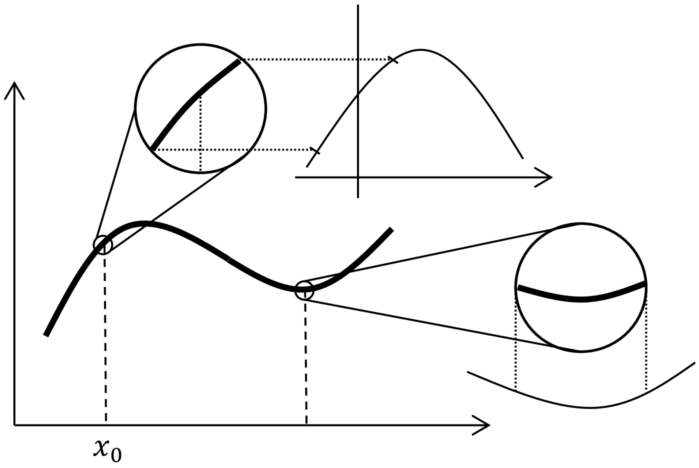
$$df = \sum_{i=1}^{m} \frac{\partial f}{\partial x_i} dx_i$$

das heißt, Funktionsänderungen können lokal durch lineare Funktionen approximiert werden.

Dies lässt sich durch den Satz von Taylor verallgemeinern: Funktionsänderungen sind lokal näherungsweise quadratisch.

Funktionen einer Variable

Hinreichend kleine Linienstücke sind näherungsweise Parabelstücken gleich.

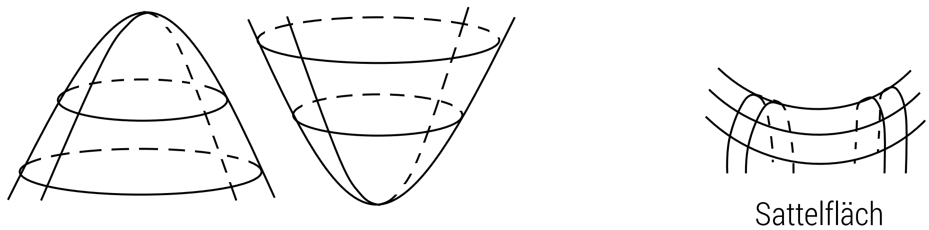


Das heißt, eine Funktion kann lokal annähernd so dargestellt werden:

$$df = f'(s)dx + \frac{f''(x)}{2}(dx)^2$$

Funktionen zweier Variablen

Flächenstücke können lokal durch Stücke von Paraboloiden dargestellt werden:



$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{1}{2}\left(\frac{\partial^2 f}{\partial x^2}dx^2 + \frac{\partial^2 f}{\partial x \partial y}dxdy + \frac{\partial^2 f}{\partial y^2}dy^2\right)$$

Allgemeine Funktionen

$$df = \nabla f \cdot d\vec{x} + \frac{1}{2} d\vec{x} H_f d\vec{x}$$

Hierbei sind:

$$\nabla f = Gradient\ von\ f$$

$$= \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_m} \end{pmatrix}$$

Also eine Folge partieller Ableitungen

$$d\vec{x} = \begin{pmatrix} dx_1 \\ \vdots \\ dx_m \end{pmatrix}$$

$H_j =$ Hesse-Matrix (Die Tabelle der zweiten partiellen Ableitungen.)

$$= \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \cdots & \cdots \\ \vdots & \ddots & \\ \vdots & & \ddots \end{pmatrix}$$

Benötigte Konzepte aus der linearen Algebra:

Skalarprodukt:

$\vec{x} \cdot \vec{y} =$ Summe der Komponentenprodukte

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = x_1 \cdot y_1 + x_2 \cdot y_2 + \ldots + x_n \cdot y_m$$

Matrixmultiplikation:

Eine Matrix wird mit einem Vektor Multipliziert ($A \cdot \vec{x}$), indem der Vektor mit jedem Zeilenvektor der Matrix (A_i) skalarmultipliziert wird. Der resultierende Vektor ist das Ergebnis der Matrixmultiplikation.

$A \cdot \vec{x}$	$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$
A_1	$A_1 \cdot \vec{x}$
A_2	$A_2 \cdot \vec{x}$
$\vdots$	$\vdots$
A_n	$A_n \cdot \vec{x}$

Optimierung mit Bedingungen zweiter Ordnung:

Untersuchen nun, ob ein gegebenes x ein Maximum ist:

$$df = \nabla f \cdot d\vec{x} + \frac{1}{2} d\vec{x} H_f d\vec{x}$$

falls $\nabla f \neq 0$ → kein Maximum, da Bedingung erster Ordnung nicht erfüllt ist.

sei $\nabla f = 0,\ x$ ist also ein kritischer Punkt.

➔ $df = \frac{1}{2} d\vec{x} H_f d\vec{x}$

falls $d\vec{x} H_j d\vec{x} > 0$ für alle $d\vec{x}$

➔ Minimum. H_f heißt dann positiv definit.

falls $d\vec{x} H_j d\vec{x} < 0$ für alle $d\vec{x}$.

➔ Maximum. H_f heißt dann negativ definit.

Trifft keines von beiden zu, heißt H_f indefinit. Ohne Weiteres lässt sich damit keine Aussage über Optimalität treffen.

Vorgehen:

Gegeben: $\max\ f(\vec{x})$

1. Kritische Punkte bestimmen

$$\frac{\partial f}{\partial x_i} \stackrel{!}{=} 0$$

→ $\vec{x}_1^*\ ,\ \vec{x}_2^*\ , \dots ,\ \vec{x}_k^*$

2. Hesse-Matrix aufstellen

$$H_j = \begin{pmatrix} \frac{\partial^2 f}{\partial {x_1}^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_m} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \ddots & \cdots & \vdots \\ \vdots & \cdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_m} \end{pmatrix}$$

3. Für jeden kritischen Punkt $\vec{x}_i^*$

1) Hesse-Matrix an $\vec{x}_i^*$ auswerten

→ $H_f\ (\vec{x}_i^*)$

2) Hauptminoren berechnen $H_1, \dots, H_m$

$$\begin{array}{cc} & \begin{array}{cccc} H_1 & H_2 & H_3 & H_4 \end{array} \\ H_f\ (\vec{x}_i\ ) = & \begin{pmatrix} \circ & \circ & \circ & \circ \\ \circ & \circ & \circ & \circ \\ \circ & \circ & \circ & \circ \\ \circ & \circ & \circ & \circ \end{pmatrix} \end{array}$$

Ein Hauptminor ist die Determinante der entsprechenden Untermatrize

3) Positiv bzw. negativ Def. anhand der Vorzeichen (VZ) der Hauptminoren bestimmen

HM	H_1	H_2	$\dots$	H_n
VZ	$+$	$+$	$\dots$	$+ \ \rightarrow neg.Def$
	$-$	$+$	$\dots$	$+/-$ $\rightarrow neg.Def$

Alternierend mit Minus zuerst

Pos. Def. $\vec{x}H_j\vec{x} > 0 => minimum$

Neg. Def. $\vec{x}H_j\vec{x} < 0 => maximum$

Anmerkung: Konvexe Funktionen sind überall pos. def.
Konkave Funktion sind überall neg. def.
Konkove/Konvexe Funktionen haben auch nur ein einziges Maximum bzw. Minimum

Optimierung mit Gleichheitsnebenbedingung

Allgemeine Form:

$$\max \overbrace{f(\vec{x})}^{ZF}$$

$\vec{y}(\vec{x}) = \vec{c}$] NB (zulässiger Bereich)

EV

Ausgeschrieben:

$$\max f(x_1, \dots, x_n)$$

$$g_1(x_1, \dots, x_n) = c_1$$
$$\vdots$$
$$g_m(x_1, \dots, x_n) = c_m$$

Bsp. 1:

$$\max f(x_1, x_2)$$

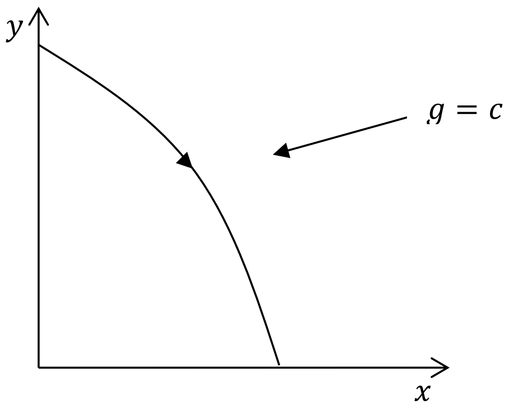
$$g(x_1, x_2) = c$$

Zulässige Punkte:

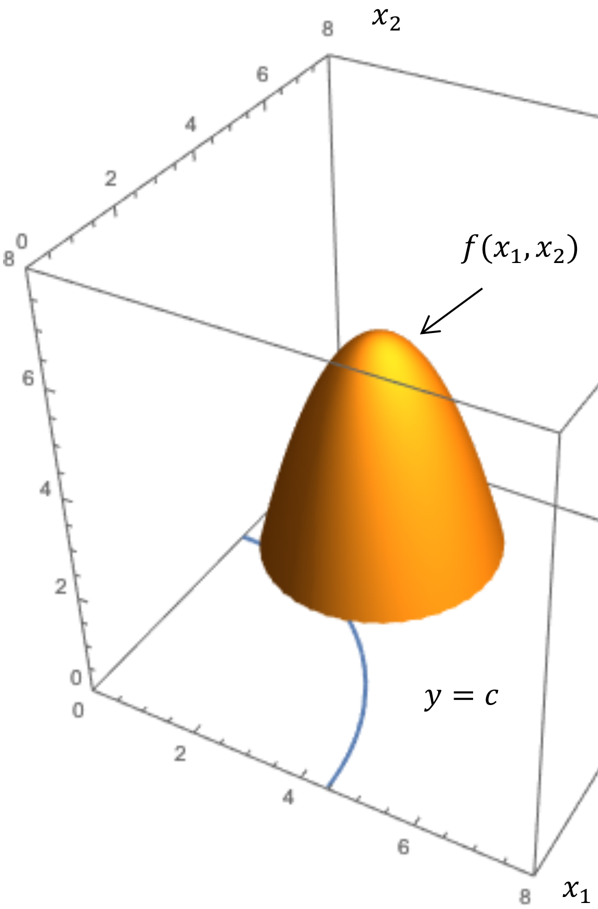
$g_1(x_1, x_2) = c_1$ 2 Variablen, 1 Gleichung

Lösung ist also 2 – 1 = 1 dimensional (lokal als Funktion in einer Variable darstellbar)

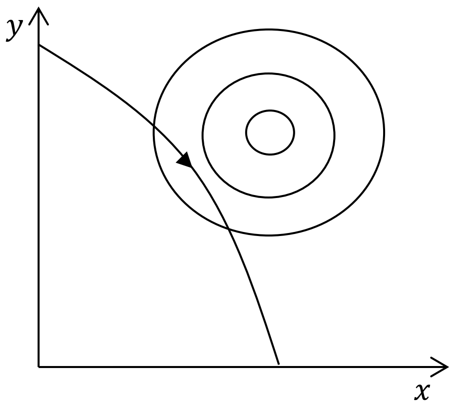
D.h. der zulässige Bereich ist eine Linie. Die Zielfunktion ZF ordnet jedem Punkt auf dieser Linie $g = c$ einen Wert zu.



Zielfunktion



Niveaulinien der Zielfunktion

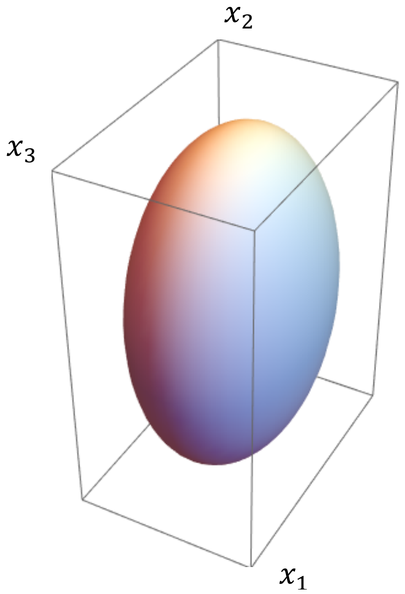


Bsp. 2:

$\max f(x_1, x_2, x_2)$

$g(x_1, x_2, x_2) = c$ 3 Variablen, 1 Gleichung

Lösung ist 3-1 = 2 dimensional (lokal als Funktion in 2 Variablen darstellbar)



Niveaumengen $f(x_1, x_2, x_3) = c_1, c_2, \dots$ der Zielfunktion, welche hier nicht abgebildet sind, sind auch Flächen im Raum.

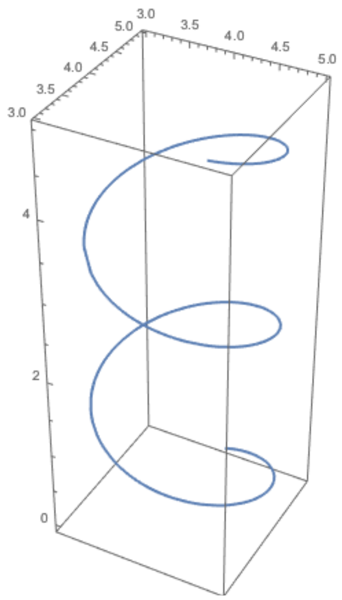
Bsp. 3:

$\max f(x_1, x_2, x_3)$

s.t. $g_1(x_1, x_2, x_3) = c_1$ 3 Variablen, 2 Gleichungen

$g_2(x_1, x_2, x_3) = c_2$ 3 – 2 = 1 Dimensional

→ Linie



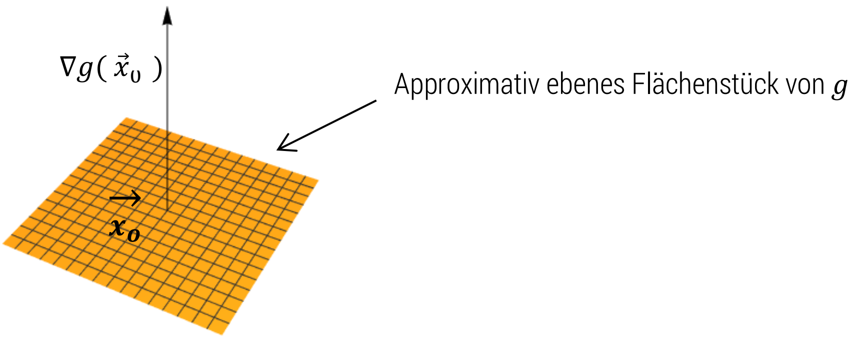
Satz 1 (Bedeutung des Gradienten):

Gegeben sei eine (Hyper-)Fläche

$g(\vec{x}) = c\,, \qquad \vec{x} \in \mathbb{R}^{\boldsymbol{n}}$

Dann $\nabla y\left(\vec{x}_0\right) \perp g \ in\ x_0$

Auf Deutsch: Der Gradient einer Fläche steht senkrecht auf einem approximativ ebenen Teilstück dieser Fläche.



Vorbemerkungen:

- Eine Ebene lässt sich durch folgende Gleichung darstellen (sog. implizite Darstellung)

  $\vec{n} \cdot (\vec{x} - \vec{x}_0) = 0$

  Die Ebene besteht aus allen Punkten $\vec{x}$ die diese Gleichung erfüllen, wobei $\vec{x}_0$ ein Stützpunkt ist.

- Bedeutung des Skalarproduktes

  $\vec{a} \cdot \vec{b} = |\overrightarrow{a}| \cdot |\vec{b}| \cdot \cos\ (\measuredangle(\vec{a}, \vec{b}))$

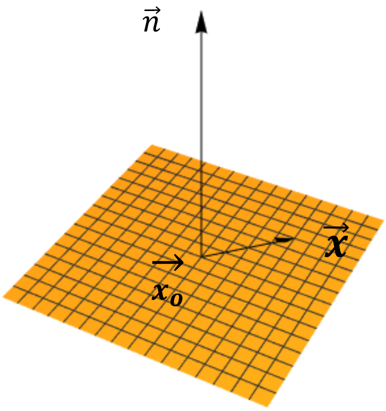
  d.h. $\quad \vec{a} \perp \vec{b} :\Leftrightarrow \measuredangle(\vec{a}, \vec{b}) = {}^{+}/{-}\ \frac{\pi}{2}$ d.h. der Winkel zwischen $\vec{a}$ und $\vec{b}$ ist 90 Grad.

  $$\Leftrightarrow \cos\left(\measuredangle(\vec{a}, \vec{b})\right) = 0$$

  $$\Leftrightarrow \vec{a} \cdot \vec{b} = 0$$

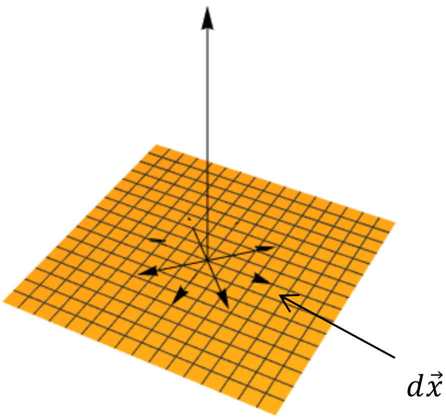
  Skalarprodukt zweier Vektor = 0 ⇔ Vektoren sind senkrecht zueinander

- Das $\vec{n}$ aus der obigen Gleichung ist der sog. Normalenvektor. Nach der vorherigen Bemerkung steht er senkrecht auf allen Vektoren in der Fläche.



- Ein Vektor $\vec{v}$ steht senkrecht auf einer Fläche $\vec{g} = \vec{c}$ in $\vec{x}_0$ wenn annährend gilt, dass $\vec{v} \perp d\vec{x} = \vec{x}_1 - \vec{x}_0$ und $\vec{x}_1, \vec{x}_0$ in $\vec{g} = \vec{c}$

  Bsp.: Steht senkrecht auf allen $d\vec{x}$



Beweis:

Linearisieren nun die im Allgemeinen krumme Fläche g

Seien $\vec{x}_0$, $\vec{x}_1 = \vec{x}_0 + d\vec{x}$ beide in g

d.h. $g(\vec{x}_0) = c$ und $g(\vec{x}_1) = c$

=> $g(\vec{x}_1) - g(\vec{x}_0) = 0$

=> $dg = 0$

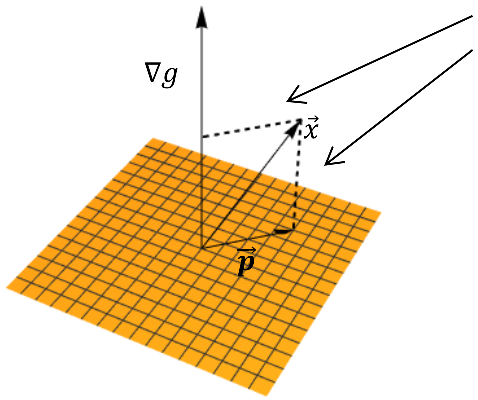
=> $\nabla g \cdot \underbrace{dx}_{(\vec{x}_1 \; - \vec{x}_0 \;)} = 0$ (siehe Abschnitt zu Differentialen)

d.h. für hinreichend nahe $\vec{x}_1$ an $\vec{x}_0$ ist die Fläche approximativ eben und ∇g steht senkrecht auf ihr.

<u>Satz 2: („Darstellungssatz")</u>

- Gegeben sei eine Fläche $g(\vec{x}) = c$, dann lässt sich jeder Punkt $\vec{p}$ im Raum darstellen als:

  $\vec{x} = \alpha_1 \nabla g + \alpha_2 \vec{p}$ mit $\vec{p} = \vec{p}_1 - \vec{p}_0$, und $\vec{p}_1, \vec{p}_0$ aus $g = c$ und $\vec{p} \perp \nabla g$



Durch Projektion von $\vec{x}$ auf $g = c$ und ∇g

In Worten: Die Fläche und ihr Gradient bilden lokal eine Art Koordinatensystem.

<u>Satz 3: (Zuwachs im Gradientenhalbraum)</u>

- Gegeben ein Punkt $\vec{x}_0$. Sei $h := f(\vec{x}_0)$ der Funktionswert (auch Niveau genannt) in $\vec{x}_0$ und $h = f(\vec{x})$ die Niveaumenge zu $\vec{x}_0$ (d.h. alle $\vec{x}$, die das gleiche Niveau wie $\vec{x}_0$ haben).

  Dann gilt $df > 0$ für $\forall\, d\vec{x} = \alpha_1 \nabla f + \alpha_2 \vec{p}$, und $(\nabla f \cdot \nabla \vec{p}) = 0\; \alpha_1 > 0$

  Auf Deutsch: Die Funktion steigt, wenn wir uns in Richtung des Gradienten bewegen, d.h. in den Halbraum, in den der Gradient hineinzeigt.

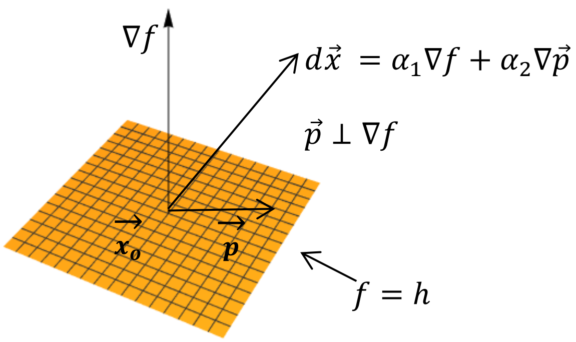
Beweis:

$$df = \nabla f \cdot d\vec{x} = \nabla f \cdot (\alpha_1 \nabla f + \alpha_2 \nabla \vec{p}\,)$$

$$= \alpha_1(\nabla f \cdot \nabla f) + \alpha_2(\nabla f \cdot \nabla \vec{p})$$

$$= \alpha_1(\nabla f \cdot \nabla f) > 0$$

Bildlich:



$d\vec{x}$ ragt in den Halbraum hinein, in welchen der Gradient zeigt.

Satz 4: (Orthogonalität des Gradienten der Zielfunktion auf dem zulässigen Bereich im Optimum)

- Im Optimum von $\begin{array}{l} max/min\, f(\vec{x}) \\ \vec{g}(\vec{x}) = \vec{c} \end{array}$

  steht ∇f senkrecht auf $\vec{g} = \vec{c}$

  Beweis: (Per Widerspruch)

  Sei x_0 ein Maximum, $f(x_0) := h$ und $h = f(\vec{x})$ die Niveaumenge des Optimums. Es gelte nun das Gegenteil des behaupteten Satzes, d.h. ∇f steht nicht auf allen $d\vec{x}$ senkrech.

  Sei $d\vec{x}$ mit $\vec{x}_1$, $\vec{x}_0$ aus $\vec{g} = \vec{c}$, s. d.

  $$d\vec{x} \not\perp \nabla f \quad \text{d.h.} \quad d\vec{x} \cdot \nabla f \neq 0$$

Stellen nun $d\vec{x}$ per „Darstellungssatz" dar durch

$d\vec{x} = \alpha_1 \nabla f + \alpha_2 \nabla \vec{p}$ wobei $\vec{p}$ in $h = f$ liegt.

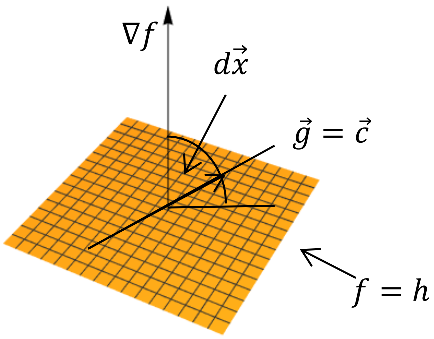
$\alpha_1 \neq 0$, denn $d\vec{x} \cdot \nabla f = \alpha_1 \underbrace{\nabla f \cdot \nabla f}_{>0} \neq 0$ (aus der Widerspruchsannahme).

(Zuwachs im Gradientenhalbraum)

Falls $\alpha_1 > 0$ => $df > 0$ für $d\vec{x}$, d.h. der Funktionswert kann lokal vergrößert werden, was einem Maximum widerspricht, denn $df = \nabla f \cdot d\vec{x} > 0$

Falls $\alpha_1 < 0$ => $df > 0$ für $-d\vec{x}$, analog

Bildlich:



In Worten: Wenn der Gradient nicht senkrecht auf dem zulässigen Bereich $\vec{g} = \vec{c}$ liegt, muss ein Teil des zulässigen Bereichs in den Gradientenhalbraum hineinragen. Auf diesem Teil könnten wir dann den Funktionswert weiter vergrößern.

Satz 5: (Gradientekombinationen des zulässigen Bereiches Spannen den zum zulässigen Bereich orthogonalen Raum auf)

- Gegeben sei $\vec{g}(\vec{x}) = \vec{c}$, also

$$\begin{bmatrix} g_1(\vec{x}) = c_1 \\ \dots \\ g_m(\vec{x}) = c_m \end{bmatrix}$$

Dann ist jeder senkrechte Vektor $\vec{v}$ auf $\vec{g} = \vec{c}$ eine Gradientenkombination

$$\lambda_1 \nabla g_1 + \dots + \lambda_m \nabla g_m$$

- D.h. die Gradienten der Nebenbedingungen spannen den orthogonalen Raum zum zulässigen Bereich auf. Anders gesagt, sie generieren alle orthogonalen Vektoren

  Beweis:

  Gegeben sei $\vec{x}_0$ aus $\vec{g} = \vec{c}$ und $\vec{x}_1$, $\vec{x}_0 + d\vec{x}$ aus $\vec{g} = \vec{c}$

  Lineariseren $\vec{g} = \vec{c}$

$$\vec{g}(\vec{x}_1) - \vec{g}(\vec{x}_0) = 0$$

$$\Leftrightarrow \begin{bmatrix} g_1(\vec{x}_1) - g_1(\vec{x}_0) = 0 \\ \dots \\ g_m(\vec{x}_1) - g_m(\vec{x}_0) = 0 \end{bmatrix}$$

$$\Leftrightarrow \begin{bmatrix} \nabla g_1 \cdot d\vec{x} = 0 \\ \nabla g_2 \cdot d\vec{x} = 0 \\ \dots \\ \nabla g_m \cdot d\vec{x} = 0 \end{bmatrix}$$

d.h. $\vec{g} = \vec{c}$ ist lokal der Schnitt von m Ebenen

Dim $\underset{\text{Var.}}{\underset{\uparrow}{n}} - \underset{\text{Gl.}}{\underset{\uparrow}{m}}$

- Sehen auch sofort, dass $\nabla g_i \perp d\vec{x}$ $\quad i = 1, \dots, m$

  d.h. die Gradienten stehen alle senkrecht auf allen $d\vec{x}$.

- Ferner, wenn ∇g_i linear unabhängig sind, bilden diese einen m dimensionalen Raum, der senkrecht auf $\vec{g} = \vec{c}$ in $\vec{x}_0$ steht

  d.h. $\vec{v} = \sum \lambda_i \nabla g_i \perp d\vec{x} = \vec{x}_1 - \vec{x}_0$, $\vec{x}_1$, $\vec{x}_0$ aus $\vec{g} = \vec{c}$

  für alle möglichen λ_i

- Die Dimension von $\vec{g} = \vec{c}$ in $\vec{x}_0$ ist aber $\overset{\text{Var.}}{\overset{\downarrow}{n}} - \overset{\text{Gl.}}{\overset{\downarrow}{m}}$

- D.h. der Raum der Gradientenkombinationen und der Unterraum $\vec{g} = \vec{c}$ haben zusammen die Dimension n, messen also ganz $\mathbb{R}^n$ aus und stehen senkrecht zueinander.

- Jeder Punkt $\vec{v}$ aus $\mathbb{R}^n$ lässt sich dann darstellen als $\vec{v} = \vec{a} + \vec{b}$ und $\vec{a} \perp \vec{b}$ wobei $\vec{a}$ eine Gradientenkombination ist und $\vec{b}$ der lineare Unterraum von $\vec{g} = \vec{c}$ in $\vec{x}_0$

Daher gilt: $\vec{v} = \overrightarrow{(a + b)} \perp \vec{g} = \vec{c}$

$\Leftrightarrow \forall d\vec{x} \qquad \vec{a} + \vec{b} \perp d\vec{x}$

$\Leftrightarrow \forall d\vec{x} \left((\vec{a} + \vec{b}) \cdot d\vec{x}\right) = 0$

$\Leftrightarrow \forall d\vec{x} \ \underbrace{\vec{a} \cdot d\vec{x}}_{=0} + \vec{b} \cdot d\vec{x} = 0$

$\Leftrightarrow \forall d\vec{x} \qquad \vec{b} \cdot d\vec{x} = 0$

Insbesondere für $d\vec{x} = \vec{b}$ also $||\vec{b}||^2 = 0$, damit auch $\vec{b} = 0$

Also $\vec{v} = \vec{a}$ und $\vec{a}$ ist eine Gradientenkombination, falls $\vec{v}$ senkrecht auf $\vec{g} = \vec{c}$ steht

$$\vec{v} = \lambda_1 \nabla g_1 + \ldots + \lambda_n \nabla g_m$$

<u>Satz 6: (Lagrange-Verfahren)</u>

Jede Lösung von $\underset{\text{s.d } \vec{g}(\vec{x}) = \vec{c}}{max/min\, f(\vec{x})}$

erfüllt $\nabla_{\vec{x}} L(\vec{x}, \vec{\lambda}) = 0$, $\nabla_{\vec{\lambda}} L(\vec{x}, \vec{\lambda}) = 0$

mit $L(\vec{x}, \vec{\lambda}) = f(\vec{x}) + \vec{\lambda} \cdot (\vec{g}(\vec{x}) - \vec{c})$

Anders: $\nabla f = \sum \lambda_i \nabla g_i$
$\vec{g}(\vec{x}) = \vec{c}$

<u>Beweis:</u>

Sei $\vec{x}_0$ ein Optimum, dann folgt:

Aus Satz 4 **=>** $\nabla f \perp (\vec{g} = \vec{c})$

Aus Satz 5 **=>** ∇f ist eine Gradientenkombination, d.h.

$$\nabla f = \lambda_1 \nabla g_i + \ldots + \lambda_m \nabla g_i$$

Im Optimum $\vec{x}_0$ also

$$\nabla f = \sum \lambda_i \nabla g_i$$

Aber nur n Gleichungen bei $m+n$ Variablen. Brauchen also noch Gleichungen für ein vollständig bestimmtes System.

Nehmen Nebenbedingung dazu für vollständig bestimmtes System:

$$\left.\begin{array}{l}\left.\begin{array}{l}\frac{\partial f}{\partial x_1} = \sum \lambda_i \frac{\partial g_i}{\partial x_1} \\ \frac{\partial f}{\partial x_2} = \sum \lambda_i \frac{\partial g_i}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} = \sum \lambda_i \frac{\partial g_i}{\partial x_n}\end{array}\right] \text{n GL} \\ \\ \left.\begin{array}{l} g_1 = c_1 \\ g_2 = c_2 \\ \vdots \\ g_m = c_m \end{array}\right] \text{m GL}\end{array}\right] \begin{array}{l} n+m \text{ GL bestimmen} \\ n+m \text{ Var vollständig}\end{array}$$

FOC also:

$$\nabla f = \sum \lambda_i \nabla g_i \ , \ \vec{g}(\vec{x}) = \vec{c}$$

Oder kürzer:

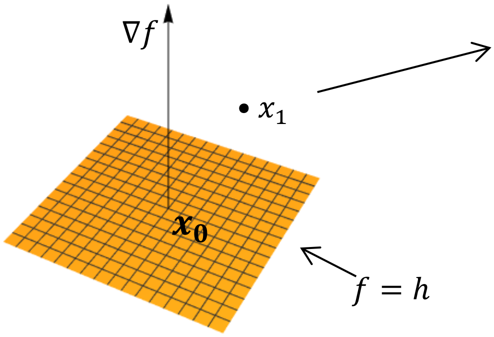
$$L(\vec{x}, \vec{\lambda}) = f - \sum \lambda_i \nabla g_i$$

$$\nabla_{\vec{x}} L \overset{!}{=} 0 \leftrightarrow \nabla f - \sum \lambda_i \nabla g_i = 0$$

$$\nabla_{\vec{\lambda}} L \overset{!}{=} 0 \leftrightarrow \vec{g}(\vec{x}) - \vec{c} = 0$$

Zusammenfassung: (alles lokal)

- f steigt in Richtung des Gradienten

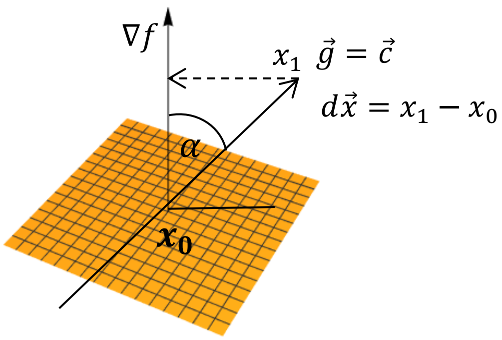


$$f(x_1) - f(x_0) > 0$$
$$\|$$
$$df = \nabla f \cdot d\vec{x}$$

Im Maximum

- $\nabla f \perp \vec{g} = \vec{c}$, ∇f ist orthogonal zu $d\vec{x}$ in $\vec{x}_0$

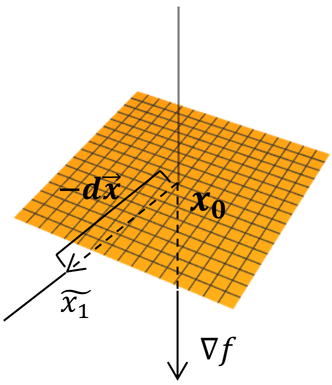
Falls nicht



$$f(x_1) - f(x_0) > 0$$
$$\underset{x_0}{\rightarrow}$$

→ kein Maximum

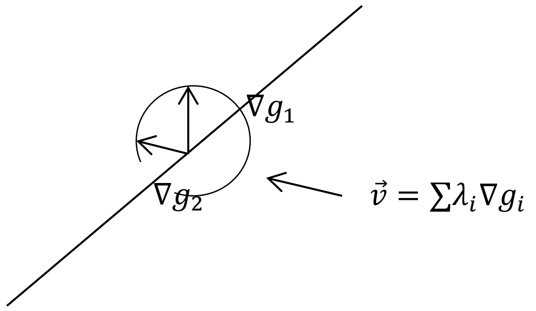
Oder



$f(\widetilde{x_1}) - f(x_0) < 0$

$\nabla f \not\perp \vec{g} = \vec{c}$ also kein Max/Min

- Gradientenkombinationen spannen den Raum auf

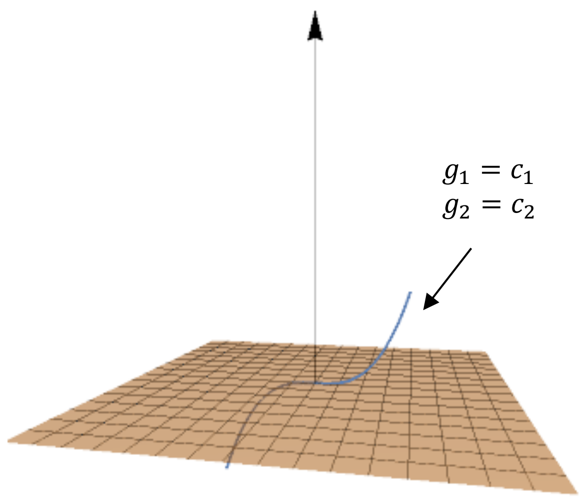


<u>Bedingungen zweiter Ordnung</u>

(engl. <u>S</u>econd <u>O</u>rder <u>C</u>onditions, SOC)

Die Bedingungen erster Ordnung (eng. first order conditions/FOC) sind nur notwendig aber nicht hinreichend.

Im folgenden Beispiel ist der Ursprung kein Optimum, erfüllt aber dennoch die Bedingungen erster Ordnung.



$$\left.\begin{matrix} g_1 = c_1 \\ g_2 = c_2 \end{matrix}\right\} \begin{bmatrix} x_3 - x_1^3 \\ x_2 = 0 \end{bmatrix} \Leftrightarrow \begin{bmatrix} x_3 = {x_1}^3 \\ x_2 = 0 \end{bmatrix} \Leftrightarrow \begin{pmatrix} x_1 \\ 0 \\ x_1^3 \end{pmatrix}$$

$$f(x_1, x_2, x_3) = x_3$$

$$f = h = 0$$

$L = f - \lambda_1 \, g_1 - \lambda_2 g_2$

$\nabla_x L = 0 \Leftrightarrow$

$\frac{dL}{dx_1} = 0 + \lambda_1 * 3x_1^2 \overset{!}{=} 0 \Leftrightarrow 3x_1^2 = 0 \Leftrightarrow x_1 = 0$

$\frac{dL}{dx_2} = 0 + \lambda_2 \overset{!}{=} 0 \Leftrightarrow \lambda_2 = 0$

$\frac{dL}{dx_3} = 1 - \lambda_1 \overset{!}{=} 0 \Leftrightarrow \lambda_1 = 1$

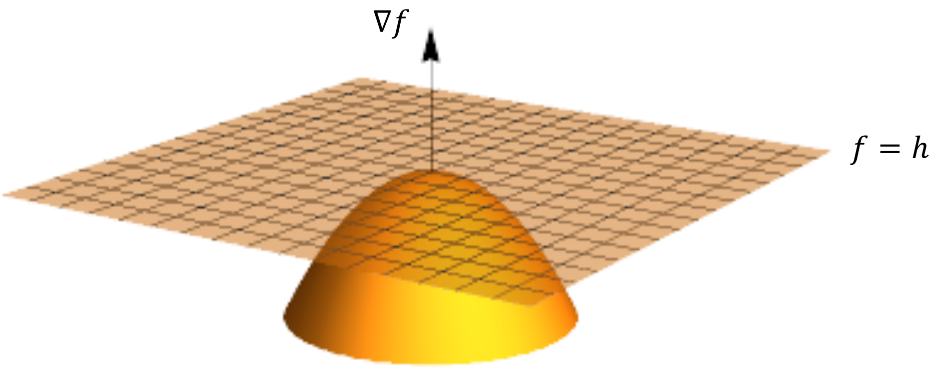
$\Leftrightarrow x_3 = 0$
$\Leftrightarrow x_2 = 0$

$f(x^*) = G$

Aber: $f\left(x^* + \begin{pmatrix} dx_1 \\ 0 \\ 0 \end{pmatrix}\right) > f(x^*)$

SOC garantieren, dass sich der zulässige Bereich von der Niveauebene vollständig auf eine Seite wegkrümmt.

D.h. der zulässige Bereich liegt lokal vollständig in einem Halbraum.



Vorgehen: mit Nebenbedingung (Für Detailliertes Vorgehen, siehe Übung)

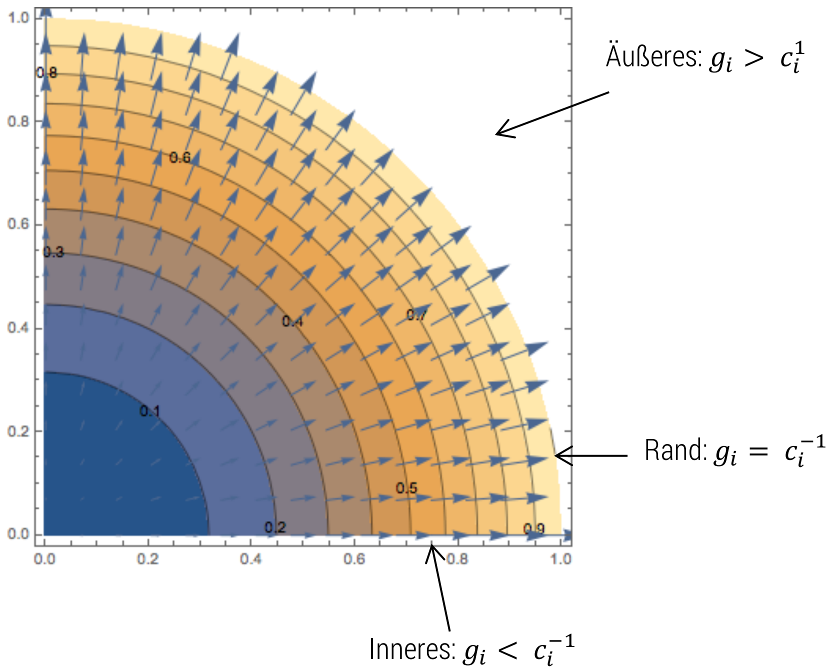
1. Lagrange Funktion aufstellen
2. Kritische Punkte finden
3. Bedingungen zweiter Ordnung anwenden:
   i. Geänderte Hesse Matrix aufstellen
   ii. Geänderte Hauptminoren berechnen beginnend mit H_{m+1}
   iii. VZ- Kriterium/ Determinantenkriterium anwenden
      a. Alternierend beginnen mit $(-1)^{m+1}$ -> max
      b. Gleich bleibend mit $(-1)^m$ -> min

Karush- Kuhn- Tucker- Verfahren (KKT)

$\max f(\vec{x})$

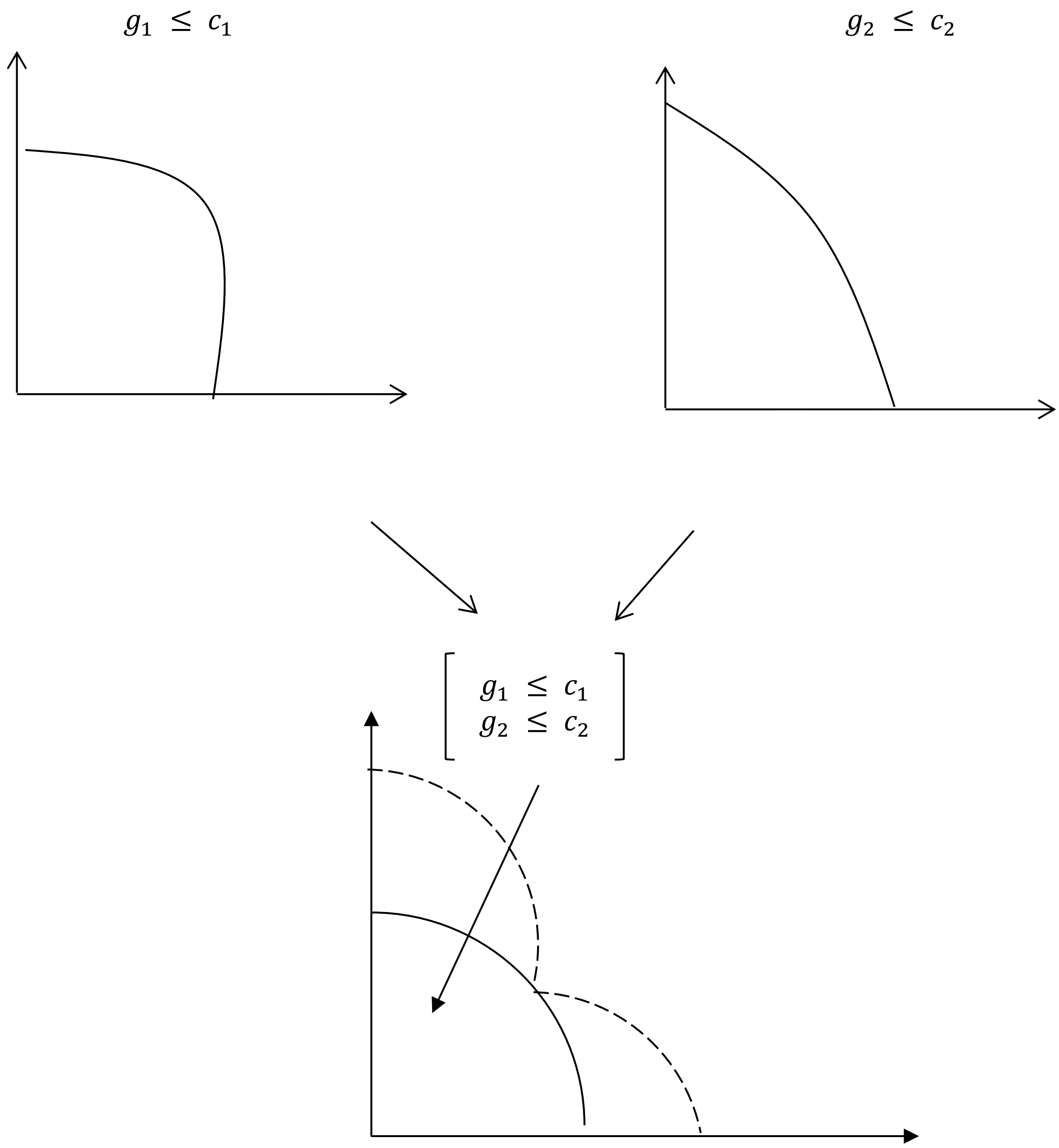
$$[\vec{g}(\vec{x}) \leq \vec{c}\,] \Leftrightarrow \begin{bmatrix} \vec{g}(\vec{x}) \leq \vec{c} \\ \vdots \\ g_m(\vec{x}) \leq c_m \end{bmatrix}$$

- Zulässiger Bereich ist ein Schnitt von m Mengen
- Wie sieht eine dieser Mengen i aus?



- Zulässiger Bereich alle x-Werte die g_i kleiner als c_i machen
- In jedem Punkt zeigt der Gradient in Richtung steigender Funktionswerte von g_i

- Visualisierung der Schnittoperation (simultane Gleichung)



$\Rightarrow$ Randpunkte erfüllen mindestens eine Gleichung $g_i = c_i$

Satz:

Gegeben sei ein Optimierungsproblem

$$\max f(\vec{x})$$

$$\text{s.d. } \vec{g}(\vec{x}) \leq \vec{c}$$

mit $\vec{x}$ aus $\mathbb{R}^n$ und $\vec{g}$, $\vec{c}$ aus $\mathbb{R}^m$ (m Gleichungen und n Entscheidungsvariablen),

dann erfüllt jede Lösung folgendes System von Gleichungen und Ungleichungen

$$\nabla f = \sum \lambda_i \nabla g_i$$

Mit $\lambda_i \geq 0\ u.\, g_i \leq c_i\ u.\, \lambda_i(g_i - c_i) = 0$ (komplementärer Schlupf , engl. complementary slackness),

d.h. bei einer der ersten Ungleichungen gilt die Gleichheit.

Beweis:

- Sei $\vec{x}$ eine Lösung für ein obiges Optimierungsproblem
- $\vec{x}$ ist entweder ein innerer Punkt oder ein Randpunkt

$\Rightarrow$ Falls innerer Punkt:

$\Rightarrow \vec{g}(\vec{x}) < \vec{c}$, d.h. keine der Nebenbedingungen greift

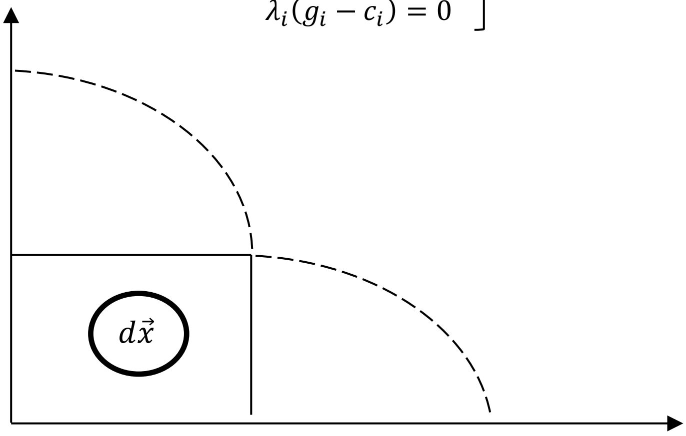
$\Rightarrow$ Können sich lokal in alle Richtungen $d\vec{x} \in \mathbb{R}^m$ im zulässigen Bereich bewegen

$\Rightarrow df = \nabla f\ d\vec{x} = 0$ für alle $d\vec{x}$ (in einem hinreichend kleinem Bereich)

$$\Leftrightarrow \nabla f = 0$$

$$\Leftrightarrow [\ \nabla f = \sum \lambda_i \nabla g_i \qquad \lambda_i = 0\ ]$$

$$\Leftrightarrow \begin{bmatrix} \nabla f = \sum \lambda_i \nabla g_i & \lambda_i \geq 0\ u.\, g_i \leq c_i \\ & \lambda_i(g_i - c_i) = 0 \end{bmatrix}$$



$\Rightarrow$ Falls Randpunkt, greifen einige NB und andere nicht. Benennen um, so dass

$\rightarrow$ Greifen: $\quad g_i = c_i \quad i = 1, \dots, k$

$\rightarrow$ Nicht Greifen: $\quad g_i < c_i \quad i = k+1, \dots, m$

$\Rightarrow$ Betrachten die Nebenbedingung nun lokal (linearisieren also)

$$g_i(\vec{x}) \leq \vec{c} = g_i(\vec{x}^*) \qquad i = 1, \dots, k$$

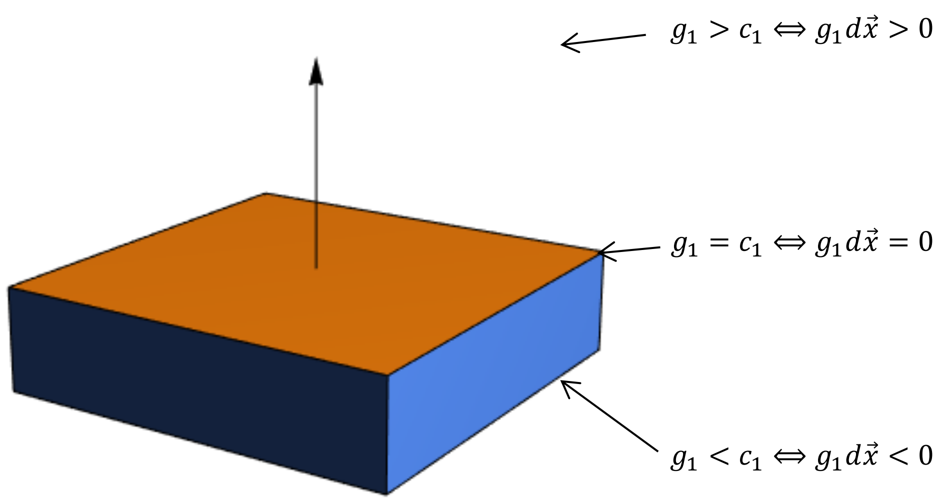
$$\Leftrightarrow g_i(\vec{x}) - \vec{c}_i \leq 0$$

$$\Leftrightarrow g_i(\vec{x}) - g_i(\overrightarrow{x^*}) \leq 0$$

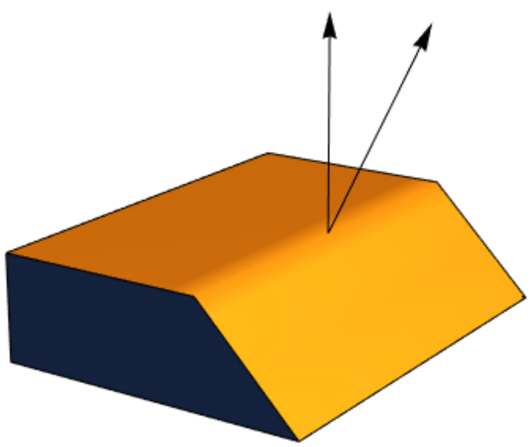
$$\Leftrightarrow dg_i \leq 0 \Leftrightarrow \nabla g_i \vec{x} \leq 0$$

Bildlich

- $g_1 \leq c_1$



Schnitt:

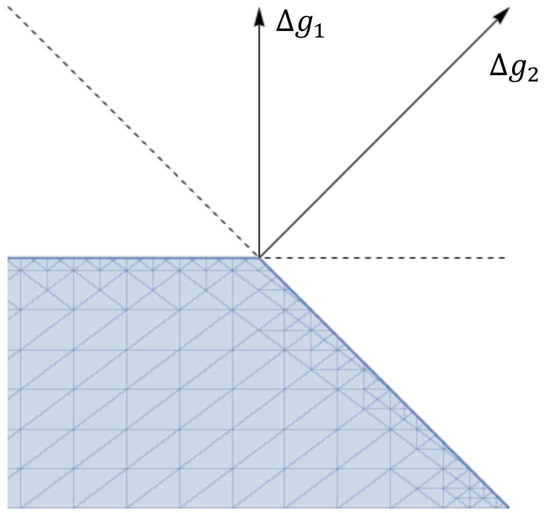


- $g_1 \leq c_1$
- $g_2 \leq c_2$

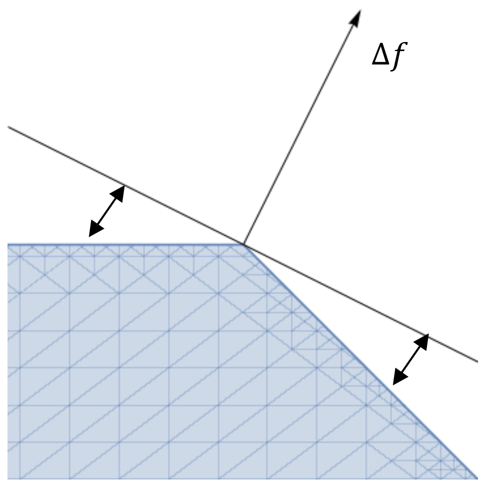
$\Delta\, g_1 d\vec{x} \leq 0$ und $\Delta\, g_2 d\vec{x} \leq 0$

lokal ist der zulässige Bereich ein Schnitt von Halbräumen

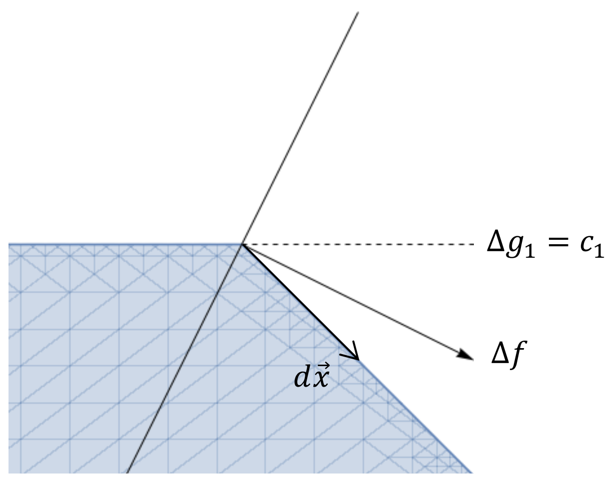
Letzteres seitlich gesehen:



Es muss folgendes gelten:



Sonst:



D.h. sonst liegt Δf im negativen Halbraum von g_1 d.h. $\nabla g_i * f_i < 0$ und wir können uns einen Schritt d$\vec{x}$ im zulässigen Bereich $\nabla g_i * d\vec{x} < 0 = 0$ bewegen.

<u>Nun allgemein</u>

Da $\mathrm{i} = 1, \dots, \mathrm{k}$ in $\vec{x}^*$ greifen so ist $\vec{x}^*$ auch eine (lokale) Lösung für

$$\max f(\vec{x})$$

$$\begin{gathered} g_1(\vec{x}) = c_1 \\ \vdots \\ g_n(\vec{x}) = c_m \end{gathered}$$

Denn lokal greifen die Nebenbedingungen $i = k+1, \dots, m$ nicht, beschränken das Problem lokal also nicht.

$$\Rightarrow f = \sum_{i=1}^{k} \lambda_i \nabla g_i$$

Zurück zum Ausgangsproblem

- Behaupten, dass $\lambda_i \geq 0$ für alle i, denn sonst sei ein $\lambda_i < 0$

Wählen nun d$\vec{x}$, so dass $\nabla g_1 d\vec{x} < 0$ (siehe letztes Bild) und

$$\begin{gathered} \nabla g_2 d\vec{x} = 0 \\ \vdots \\ \nabla g_n d\vec{x} = 0 \end{gathered}$$

Dann gilt

$f(\vec{x}^* + d\vec{x}) - f(\vec{x}^*)$

$= df = \nabla \mathrm{f}\, d\vec{x}$

$= \sum_{i=1}^{k} \lambda_i \nabla g_i d\vec{x}$

$= \underbrace{\underbrace{\lambda_i}_{<0} \underbrace{\nabla g_i d\vec{x}}_{<0}}_{>0} + \lambda_2 \underbrace{\nabla g_2 d\vec{x}}_{=0} + \dots + \lambda_m \underbrace{\nabla g_m d\vec{x}}_{=0}$

$\Rightarrow df > 0 \quad \Rightarrow$ Wäre also kein Maximum entgegen der Annahme.

Also gilt im Maximum

$\nabla f = \sum_{i=1}^{k} \lambda_i \nabla g_i$, $\lambda_i > 0$ und $g_i = c_i$, für $i = 1, \ldots, k$

$$\Leftrightarrow \nabla f = \sum_{i=1}^{k} \lambda_i \nabla g_i + \underbrace{\sum_{i=k+1}^{m} \lambda_i \nabla g_i}_{= 0}$$

$\lambda_i > 0$ und $g_i = c_i,\ \lambda_i(g_i - c_i) = 0,\ i = 1, \ldots, k$

$\lambda_i = 0$ und $g_i < c_i,\ \lambda_i(g_i - c_i) = 0,\ i = k+1, \ldots, m$

Obige Bedingungen sind ein Spezialfall von

$\Rightarrow \nabla f = \sum_{i=1}^{m} \lambda_i \nabla g_i \qquad \lambda_i \geq 0, g_i \leq c_i, \lambda_i(g_i - c_i) = 0$

<u>Vorgehen:</u>

Gegeben max $f(\vec{x})$

$$\vec{g}(\vec{x}) \leq \vec{c}$$

- Lagrange-Funktion aufstellen
- 2^m Fallunterscheidungen machen, danach ob die Nebenbedingung i greift oder $\boldsymbol{\lambda_i = 0}$ ist.

	Fall 1	Fall 2	…	Fall 2^n
$\lambda_1 \geq 0 \quad g_1 \leq c_1$	$g_1 = c_1$	$\lambda_1 = 0$		$\lambda_1 = 0$
$\lambda_2 \geq 0 \quad g_2 \leq c_2$	$g_2 = c_2$	$g_2 = c_2$		$\vdots$
$\vdots$	$\vdots$	$\vdots$		$\vdots$
$\lambda_m \geq 0 \quad g_m \leq c_m$	$g_m = c_m$	$g_m = c_m$		$\lambda_1 = 0$

$\uparrow$ Entspricht innerer Lösung

- Bekommen für jeden Fall m Gleichungen aus dem komplementären Schlupf und zusätzlich noch n Gleichungen aus der Gradientengleichung.
  Haben insgesamt also n+m Gleichungen und n+m Variablen, womit das Gleichungssystem als bestimmt gedacht werden kann.

- Lösen das Gleichungssystem für jeden Fall und prüfen ob die Lösung den komplementären Schlupf erfüllt.

- Falls ja haben wir eine Lösung.

6. Homogene Differentialgleichungen

6.1 Harrod-Domar-Model

Annahmen:

- Makroökonomische Produktionsfunktion

  $Y = f(K)$

  Speziell: constant returns to scale

  $Y = c \cdot K \qquad , \qquad c > 0$

  Wobei c die konstante Grenzproduktivität des Kapitals darstellt und $f(0) = 0$.

- konstante Sparquote s

  $S = sY \qquad , \qquad s > 0$

- Kapitalakkumulation:

  $\dot{K} = I$

- Makroökonomisches Gleichgewicht

  $I = S$

- positiver Anfangsbestand $K(0) > 0$

Bezeichnungen:

Y : makroökonomischer Output

K : Kapitalbestand

I : Investitionen

S : Sparen

c : Grenzproduktivität des Kapitals

s : Sparquote

Berechne die gleichgewichtige Wachstumsrate des Outputs:

$$g_Y := \frac{\dot{Y}}{Y}$$

Modell

$$\begin{cases} Y = c \cdot K & \rightarrow \quad \dot{Y} = c \cdot \dot{K} \\ \dot{K} = s \cdot Y & \\ K(0) > 0 & \end{cases}$$

$$\Leftrightarrow g_Y := \frac{\dot{Y}}{Y} = \frac{c \cdot \dot{K}}{c \cdot K} = \frac{\dot{K}}{K}$$

$$\frac{\dot{K}}{K} = \frac{s \cdot Y}{K} = sc$$

$$\rightarrow g_Y = g_K = sc$$

Der Output wächst mit der Sparquote und der marginalen Kapitalproduktivität

Steady-State:

Def. Alle Wachstumsraten müssen konstant, nicht notwendigerweise gleich sein,

hier erfüllt, da sc konstant ist

Zur Berechnung des Pfads $K(t)$ ist zu lösen:

Differentialgleichung:

$$\dot{K} = scK$$

$$\Leftrightarrow \frac{\dot{K}}{K} = sc$$

Unbestimmte Integration beider Seiten nach t :

$$\int \frac{\dot{K}}{K} dt = \int sc\, dt$$

$$\Leftrightarrow \int \frac{1}{K}\frac{dK}{dt} dt = sc \int dt$$

$$\Leftrightarrow \int \frac{1}{K} dt = sc(t + C_1)$$

$$\Leftrightarrow \ln|K| + C_2 = sct + scC_1$$

$$\Leftrightarrow e^{\ln K} \cdot e^{C_2} = e^{sct} \cdot e^{scC_1}$$

$$K(t) = e^{sct} \cdot C$$

Mit Integrationskonstante $C := e^{scC_1 - C_2}$

Berechnung von C :

$$K(0) = e^{sc0} \cdot C$$

$$= C$$

bzw. $C = K(0)$

6.2 Lösung von Differentialgleichungen

Eine homogene Differentialgleichung erster Ordnung ist von der Form:

$$\dot{x}(t) = f(x(t))$$

Beispiel einer homogenen Differentialgleichung 1. Ordnung:

$$\dot{x}(t) + ax(t) = 0 \qquad a \in \mathbb{R} \text{ (konstanter Parameter)}$$

$$x(0) = x_0 > 0$$

Die boundary condition $x(0) = x_0$ dient zur Spezifikation der Integrationskonstanten.

Beispiel einer inhomogenen Differentialgleichung 1. Ordnung:

$$\dot{x}(t) + ax(t) = b \qquad a \in \mathbb{R}, b > 0$$

Charakteristisch für Differentialgleichungen: Die Variable x kommt gleichzeitig als Ableitung nach der Zeit und als Variable in der Gleichung vor.

Lösung der Differentialgleichung:

$$\dot{x}(t) = -ax(t) \Leftrightarrow \frac{\dot{x}(t)}{x(t)} = -a$$

$$\Leftrightarrow \int \frac{\dot{x}(t)}{x(t)} dt = -\int a dt \quad \text{(unbestimmte Integration)}$$

$$\Leftrightarrow \int \frac{1}{x(t)} \frac{dx}{dt} dt = -a \int 1 dt$$

$$\Leftrightarrow \ln|x(t)| + C_0 = -at + C_1$$

Isolierung der Integrationskonstanten (C_0, C_1)

$$\Leftrightarrow e^{\ln|x(t)| + C_0} = e^{-at + C_1}$$

$$\Leftrightarrow x(t)e^{C_0} = e^{-at + C_1}$$

$$\Leftrightarrow x(t) = e^{-at} \cdot \underbrace{e^{C_1 - C_0}}_{=:C}$$

$$\Leftrightarrow x(t) = e^{-at} \cdot C,$$

Bestimmung der Integrationskonstanten C über $t = 0$ mit $x(0) = C$:

$\Leftrightarrow C = x_0$, da $x(0) = x_0$ nach Annahme

Lösung: $\Leftrightarrow x(t) = x_0 \cdot e^{-at}$ (x_0: Anfangswert, a: Wachstumsrate)

7. Inhomogene Differentialgleichungen

7.1 Solow-Wachstums-Modell

Annahmen:

Output

$Y = F(K, L)$

K : Kapitalinput

L : Arbeitsinput

F : Produktionsfunktion, konstante Skalenerträge

$F(\lambda K, \lambda L) = \lambda F(K, L) \quad , \quad \lambda > 0$

Dann gilt:

$Y/L = F(K, L)/L = F(K/L, 1)$

und mit $y := Y/L \quad , \quad k := \frac{K}{L}$

$y = f(k)$, wobei $f(k) := F(k, 1)$

Sei speziell:

$f(k) = k^{\alpha} \quad , \quad 0 < \alpha < 1$

Sparen:

$\dot{k} = sf(k) \quad , \quad s > 0$

bei konstanter Bevölkerung und ohne Abschreibungen

Bevölkerungswachstum:

$\frac{\dot{L}}{L} = n \quad , \quad n$: Wachstumsrate der Bevölkerung (konstant)

$$\dot{k} = \frac{\dot{K}}{L} - \frac{\dot{L}K}{L^2} = \frac{\dot{K}}{L} - \frac{\dot{L}}{L}\frac{K}{L} = \frac{\dot{K}}{L} - nk$$

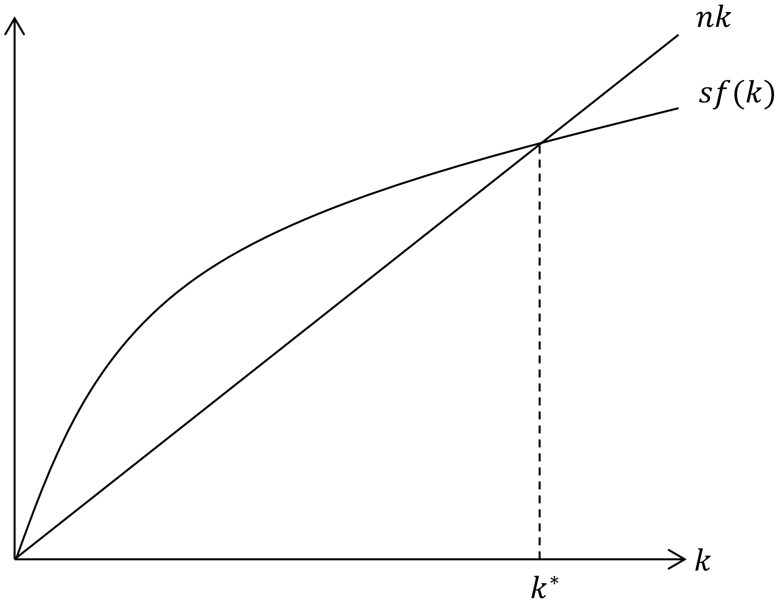
$= s \cdot Y/L - nk$

$= s \cdot f(k) - nk$

Es gibt einen Punkt der Kapitalakkumulation, bei dem der Pro-Kopf-Output nicht mehr wächst:

$\dot{k} = 0 \qquad \Leftrightarrow \qquad sf(k) = nk$

Steady-state value k^* :



Output: $y^* = f(k^*)$

<u>Wachstumsraten:</u>

$$\frac{\dot{y}}{y} = \frac{\alpha k^{\alpha-1}\dot{k}}{k^{\alpha}} = \alpha \cdot \frac{\dot{k}}{k}$$

$$\frac{\dot{k}}{k} = s \cdot f(k)/k - n$$

$$= s \cdot k^{\alpha}/k - n$$

$$= s \cdot k^{\alpha-1} - n$$

$$= s \cdot \frac{1}{k^{1-\alpha}} - n$$

$$\Rightarrow \frac{\dot{y}}{y} = s \cdot \alpha \cdot \frac{1}{k^{\alpha-1}} - \alpha n$$

Bevölkerungswachstum kann nicht verhindern, dass Wachstum Pro-Kopf ausläuft.

7.2 Lösen von inhomogenen Differentialgleichungen

sind vom Typ:

$\dot{x}_t = ax_t + b_t$ oder ausführlich

$\dot{x}(t) = ax(t) + b(t)$

Für $b(t) = 0$ lautet die Lösung bekanntlich: $x(t) = x_0 e^{at}$

Annahme:

Die Lösung für die nicht-homogene Differentialgleichung lautet: $x(t) = c(t)e^{at}$

Dann gilt ebenso: $\dot{x}(t) = \dot{c}(t)e^{at} + c(t)ae^{at}$

Einsetzen von $\dot{x}(t)$ und $x(t)$ in die Differentialgleichung

$$\rightarrow \dot{c}(t)e^{at} + c(t)ae^{at} = ax(t) + b(t)$$

$$= ac(t)e^{at} + b(t)$$

$$\Rightarrow \dot{c}(t) = b(t)e^{-at}$$

$$\Rightarrow \int \dot{c}(t)dt = b\int e^{-at}\,dt \quad , \quad f\ddot{u}r\ b(t) = b$$

$$\Rightarrow C + c(t) = -\frac{1}{a}be^{-at}$$

$$\Rightarrow \quad x(t) = -\frac{1}{a}be^{-at} \cdot e^{at} - Ce^{at}$$

$$\Rightarrow \quad x(t) = -\frac{1}{a}b - Ce^{at}$$

Zum Zeitpunkt $t = 0$

$$x(0) = -\frac{1}{a}b - C$$

$$\Rightarrow \qquad C = x_0 + \frac{b}{a} \qquad , \qquad x(0) = x_0$$

Anfangswert von x