

Angewandte Mathematik
Numerische Optimierung (Numerik 3)

Skript zur Vorlesung

Prof. Dr. Dirk Lebiedz

Wintersemester 2022/23

Universität Ulm

„Da der Plan des gesamten Universums der vollkommenste ist und vom weisesten Schöpfer festgelegt, so geschieht nichts auf der Welt, dem nicht irgendein Verhältnis des Maximums oder Minimums zu Grunde liegt. Man wird also danach streben müssen, in jeder Art naturwissenschaftlicher Probleme die Größe zu bestimmen, die einen größten oder kleinsten Wert annimmt. Damit ist ein doppelter Weg zur Lösung naturwissenschaftlicher Fragen gegeben: Aus den Endursachen mit Hilfe der Methode der Maxima oder Minima (a posteriori) oder aus den bewirkenden Ursachen (a priori). Man soll aber besonders darauf bedacht sein, die Lösung auf beiden Wegen herzuleiten. Dann wird nicht nur die eine zur Bestätigung der anderen dienen, sondern uns mit höchster Befriedigung erfüllen.“

L. Euler in “Methodus inveniendi lineas curvas maximi minimive proprietate gaudentes sive solutio problematis isoperimetrici latissimo sensu accepti” (1744)

“If you teach a man anything, he will never learn” (G.B. Shaw)

Diese Vorlesung soll unter dem gerade in der Mathematik so wichtigen Motto “Learning by doing” stehen !

„Es ist nicht genug, den Menschen ein Spezialfach zu lehren. Dadurch wird er zwar zu einer Art benutzbarer Maschine, aber nicht zu einer vollwertigen Persönlichkeit. Es kommt darauf an, dass er ein lebendiges Gefühl dafür bekommt, was zu erstreben wert ist. Er muss einen lebendigen Sinn dafür bekommen, was schön und was moralisch gut ist. Sonst gleicht er mit seiner spezialisierten Fachkenntnis mehr einem wohlabgerichteten Hund als einem harmonisch entwickelten Geschöpf. Er muss die Motive der Menschen, deren Illusionen, deren Leiden verstehen lernen, um eine richtige Einstellung zu den einzelnen Mitmenschen und zur Gemeinschaft zu erwerben.“ (Albert Einstein, Mein Weltbild)

Die erste Version des Skripts wurde im Rahmen einer Vorlesung Nichtlineare Optimierung von Prof. Dr. Dirk Lebiedz im WS 2008/2009 an der Universität Freiburg erstellt unter Mithilfe von Markus Esenwein und Jonas Unger. Das Skript basiert teilweise auf Inhalten der Vorlesungen “Optimierung I und II” eines Kollegen der Universität Münster und enthält Teile eines Skripts von Prof. Dr. Roland Herzog (TU Chemnitz) zur Vorlesung “Grundlagen der Optimierung”. Fehler und Kommentare bitte an: dirk.lebiedz@uni-ulm.de

Stand: 23. September 2022

Inhaltsverzeichnis

Kapitel 1.	Differenzierbare beschränkte finite Optimierung	11
1	Motivation und Problemstellung	11
2	Einführung und allgemeine Konzepte der Optimierungstheorie	13
2.1	Existenz von Optimallösungen	16
2.2	Grundbegriffe der Optimierung im \mathbb{R}^n	19
2.3	Klassifikation von Optimierungsproblemen	20
2.4	Beispiele für Optimierungsprobleme	20
2.5	Konvexität und Trennungssatz im \mathbb{R}^n	22
3	Nichtlineare Optimierung	30
3.1	Notwendige Optimalitätsbedingungen	30
3.2	Hinreichende Optimalitätsbedingungen	38
3.3	Sensitivitätsanalyse	41
4	Standardproblem der Nichtlinearen Optimierung im \mathbb{R}^n	45
4.1	Notwendige Optimalitätsbedingungen	45
4.2	Notwendige Bedingungen unter LICQ	46
4.3	Notwendige Bedingungen unter MFCQ	47
4.4	Notwendige Bedingungen für konvexe Optimierungsprobleme unter Slater-Bedingungen	47
5	Lineare Optimierung als Spezialfall	49
5.1	Grundbegriffe und Konzepte	49
5.2	Dualität	52
5.3	Innere-Punkte-Verfahren für LP	55
Kapitel 2.	Numerische Verfahren für finite NLP	65
6	Unbeschränkte Optimierung	65
6.1	Optimalitätsbedingungen	65
6.2	Ein allgemeines Abstiegsverfahren	65
6.3	Eine effiziente Schrittweitenstrategie: Armijo-Regel	68
6.4	Wahl der Suchrichtungen: Gradientenverfahren	69
6.5	Wahl der Suchrichtungen: Newton-Verfahren	71
7	Beschränkte Nichtlineare Optimierung (NLP)	73
7.1	Quadratische Programme (QP)	73
7.2	Lagrange-Newton-Iteration	78
7.3	Lokales SQP-Verfahren	79
7.4	Globalisierung von SQP - Verfahren	82
7.5	Innere Punkte Verfahren (IP) für NLP	83
7.6	Zusammenfassung der Theorie und Numerik differenzierbarer beschränkter Optimierungsprobleme	83

Kapitel 3.	Nichtdifferenzierbare finite Optimierung	85
8	Lagrange-Dualität und Sattelpunkte	85
8.1	Lagrange-Dualität	85
8.2	Sattelpunkt-Interpretation	90
9	Das Subdifferential konvexer Funktionen	94

Zur 'Philosophie' universitärer Lehre (in der Mathematik)

An einer Universität geht es primär um **Bildung** (und zwar im Humboldt'schen Sinne), in einem Rahmen, den die ursprüngliche Bedeutung des Namens Universität - *universitas magistrorum et scholarium* - setzt, einer '**Gemeinschaft von Lehrenden und Lernenden**'.

Grundgedanke ist nicht in erster Linie die Stoffvermittlung, sondern das gemeinsame Erarbeiten, das Erleben von Wissenschaft als dialogischen diskursiven Prozess. Daher sollte universitäre **Lehre** wann immer möglichst **verknüpft** sein mit **aktueller Forschung** (Wilhelm von Humboldts Idee der Einheit von Forschung und Lehre - Bildung durch Wissenschaft). Das werde ich an geeigneten Stellen im Vorlesungszyklus Numerik versuchen. Es macht auch tatsächlich im Sinne der modernen Wissenschaft kaum Sinn, ein Fach in Teilfelder wie Reine und Angewandte Mathematik, Analysis, Algebra und Numerik usw. aufzuteilen. Lediglich die Methoden unterscheiden sich etwas, aber die **Mathematik bildet eine Einheit** ! Auch wenn die Numerische Mathematik in weiten Teilen eine 'Angewandte Mathematik' ist, deren Entwicklung sich aber durchaus auch entlang der Linien von Hadamard (siehe nachfolgendes Zitat) vollzieht.

Jacques Hadamard in 'The Psychology of Invention in the Mathematical Field':

'The answer appears to us before the question. Practical application is found by not looking for it, and one can say that the whole progress of civilization rests on that principle. Practical questions are most often solved by means of existing theories. It seldom happens that important mathematical researches are directly undertaken in view of a given practical use: they are **inspired by** the desire which is the common motive of every scientific work, **the desire to know and to understand.**'

Die **Rolle des Dozenten** ist **nicht nur** die des '**Beibringens**', sondern insbesondere die eines Motivators und 'Leuchtturms', der versucht, Wege zu weisen, mit einem gewissen **Überblick eine 'Landkarte einer Wissenschaft'** oder spezieller eines Fachgebietes und auf Basis eigener Erfahrung einen roten Faden in Form von Konzepten und Ideen zeigt und Diskussionspartner ist. Vergessen Sie nicht, dass die Professoren auch einmal Studierende waren. Das einzige, was ein Dozent guten Studierenden voraus hat, ist Erfahrung. Dadurch hat er einen mehr oder weniger wertvollen Überblick. Ein Dozent sollte daher insbesondere diese **eigene Erfahrung weitergeben** (kein 'totes Wissen' oder reinen Lehrbuchstoff), auch bzgl. einer Bewertung dessen, was in seinen Augen wichtig ist, Schlüsselkonzepte enthält und Potenzial für neues Denken bietet, er agiert dann als Lotse auf dem Weg der Studierenden, **Wissenschaft zu ERFAHREN** ! Das ist im Idealfall eine Art Analogie zur sokratischen Mäeutik (Hebammenkunst), eine ErkenntnisFINDUNGSunterstützung im Sinne der Geburtshilfe eines (geistigen) Kindes, das in interessierten Studierenden selbst steckt. Brouwer (1881-1966, niederländ. Mathematiker) sagte über die Mathematik: ' Sie ist **mehr ein Tun als eine Lehre**', passend dazu Hermann Hesse in *Das Glasperlenspiel*: 'Die Wahrheit wird gelebt, nicht doziert.' Wie gut sich dieser 'Leitfaden für die Wissenschaft' hier in der Lehrveranstaltung Numerische Lineare Algebra umsetzen lässt, hängt an Ihnen zu mindestens gleichen Teilen wie an mir !

Mathematische Wissenschaft und Forschung haben neben dem logisch-analytischen Denken viel mit **Kreativität, Phantasie und Ideen** zu tun. Diese Qualitäten kommen in Lehrplänen oft zu kurz, werden aber gefördert durch **Begeisterung und Hingabe** für eine Sache und **tiefe Beschäftigung** mit dieser. Es gibt jede Menge lesenswerte mehr oder weniger populärwissenschaftliche Literatur zu 'tief-liegenden mathematischen Entdeckungen', die begeistern. Insbesondere möchte ich Ihnen empfehlen

- (1) S. Singh: Fermats letzter Satz
- (2) D. O'Shea: Poincarés Vermutung
- (3) M. du Sautoy: Die Musik der Primzahlen
- (4) E. Frenkel: Liebe und Mathematik
- (5) I. Stewart: Die Macht der Symmetrie
- (6) J. Derbyshire: Prime Obsession
- (7) J. Derbyshire: Unknown Quantity.

Nicht zuletzt spielen **menschlich-persönliche und soziokulturelle Aspekte** eine zentrale Rolle auf dem Bildungsweg Universität, ganz im Sinne von Albert Einstein und Werner Heisenberg.

Aus **'Mein Weltbild' (Albert Einstein, 1953):**

Es ist nicht genug, den Menschen ein Spezialfach zu lehren. Dadurch wird er zwar zu einer Art benutzbarer Maschine, aber nicht zu einer vollwertigen Persönlichkeit. Es kommt darauf an, dass er einen lebendigen Sinn dafür bekommt, was schön und was moralisch gut ist. Sonst gleicht er mit seiner spezialisierten Fachkenntnis mehr einem wohlabgerichteten Hund als einem harmonisch entwickelten Geschöpf. Er muss die Motive der Menschen, deren Illusionen, deren Leiden verstehen lernen, um eine richtige Einstellung zu den einzelnen Mitmenschen und zur Gemeinschaft zu erwerben. Diese wertvollen Dinge werden der jungen Generation durch **persönliche Kontakte mit den Lehrenden**, nicht - oder wenigstens nicht in der Hauptsache - durch Textbücher vermittelt. Dies ist es, was Kultur in erster Linie ausmacht und erhält. Dieses habe ich im Auge, wenn ich die 'humanitas' als wichtig empfinde, nicht einfach trockenes Fachwissen auf geschichtlichem und philosophischem Gebiet. Überbetonung des Konkurrenzsystems und frühzeitiges Spezialisieren unter dem Gesichtspunkt der unmittelbaren Nützlichkeit töten den Geist, von dem alles kulturelle Leben abhängig ist. Zum Wesen einer wertvollen Erziehung gehört es ferner, dass das **selbstständige kritische Denken im jungen Menschen entwickelt wird, eine Entwicklung, die weitgehend durch Überbürdung mit Stoff gefährdet wird**. Überbürdung führt notwendig zu Oberflächlichkeit und Kulturlosigkeit. Das Lehren soll so sein, dass das Dargebotene als wertvolles Geschenk und nicht als saure Pflicht empfunden wird.

Aus dem Vorwort von **'Der Teil und das Ganze' (Werner Heisenberg, 1969):**

Wissenschaft wird von Menschen gemacht. Dieser an sich selbstverständliche Sachverhalt gerät leicht in Vergessenheit, und es mag zur Verringerung der oft beklagten **Kluft zwischen den beiden Kulturen, der geisteswissenschaftlich-künstlerischen und der technisch-naturwissenschaftlichen**, beitragen, wenn man ihn wieder ins Gedächtnis zurückruft. ...

Es geht in der Wissenschaft indirekt oft um menschliche, philosophische oder politische Probleme, und der Verfasser hofft, dass gerade daran deutlich wird, **wie wenig sich die Naturwissenschaft von diesen allgemeineren Fragen trennen lässt**.

...

Felix Klein in der Einleitung seines Werkes

'Vorlesungen über die Entwicklung der Mathematik im 19. Jarhundert':

Wie keine andere Wissenschaft ist die Mathematik ein auf wenigen Grundprinzipien nach zwingenden Gesetzen aufgerichtetes Gebilde. Der Charakter der Ausschließlichkeit, der ihre Entwicklung vor der anderer Gebäude des Geistes auszeichnet und ihr die vielgrühmte ‚Klarheit‘ verleiht, bringt es mit sich, dass sie auch die am schwersten zugängliche aller Wissenschaften ist. Denn wer in die eindringen will, muss **in sich durch eigene Arbeit die ganze Entwicklung Schritt für Schritt wiederholen**; es ist doch unmöglich, auch nur einen mathematischen Begriff zu erfassen, ohne alle die davorliegenden Begriffe und ihre Verbindungen in sich aufgenommen zu haben, die zu seiner Erschaffung führten.

On Teaching Math (Auszug, V.I. Arnold)

Es folgt ein 'Plädoyer', eine Hommage an tiefliegende Verbindungen innerhalb der Mathematik (siehe Eingangszitat von André Weil), sowohl zwischen Reiner und Angewandter Mathematik/Numerik als auch zwischen Mathematik und Naturwissenschaften, die mathematische Modelle verwenden, insbesondere Physik. Hier stellt die Numerische Mathematik eine wichtige Toolbox ! Und zwar u.a. für 'Experimente' auf der Basis von Modellen und Computersimulationen. Nehmen Sie nicht alle, teilweise etwas polemisch-verbitterten, Phrasen im Text ernst, es geht um den Gesamttenor und der Fokus liegt auf Verbindung und Vernetzung. Natürlich sind auch abstrakte mathematische Konzepte per se wichtig in der modernen Wissenschaft, aber es ist erhellend, ggf. Motivationen für diese zu kennen, die in vielen Fällen genuin in den Naturwissenschaften, insbesondere der Physik liegen.

This is an extended text of the address at the discussion on teaching of mathematics in Palais de Découverte in Paris on 7 March 1997:

Mathematics is a part of physics. Physics is an experimental science, a part of natural science. Mathematics is the part of physics where experiments are cheap. The Jacobi identity (which forces the heights of a triangle to cross at one point) is an experimental fact in the same way as that the Earth is round (that is, homeomorphic to a ball). But it can be discovered with less expense. In the middle of the twentieth century it was attempted to divide physics and mathematics. The consequences turned out to be catastrophic. Whole generations of mathematicians grew up without knowing half of their science and, of course, in total ignorance of any other sciences. They first began teaching their ugly scholastic pseudo-mathematics to their students, then to schoolchildren (forgetting Hardy's warning that ugly mathematics has no permanent place under the sun). Since scholastic mathematics that is cut off from physics is fit neither for teaching nor for application in any other science, the result was the universal hate towards mathematicians - both on the part of the poor schoolchildren (some of whom in the meantime became ministers) and of the users. The ugly building, built by undereducated mathematicians who were exhausted by their inferiority complex and who were unable to make themselves familiar with physics, reminds one of the rigorous axiomatic theory of odd numbers. Obviously, it is possible to create such a theory and make pupils admire the perfection and internal consistency of the resulting structure (in which, for example, the sum of an odd number of terms and the product of any number of factors are defined).

Jacobi noted, as mathematics' most fascinating property, that in it one and the same function controls both the presentations of a whole number as a sum of four squares and the real movement of a pendulum. If, instead of using the harmonic oscillator equation for a pendulum, you use the expression involving the actual force, you get an elliptic integral for the time as a function of the angle. Taking the inverse to get the angle as a function of the time you get an elliptic function, by Jacobi's definition as inverse functions of elliptic integrals. Jacobi made that definition in an analogy to how you get \sin , \cos and \tan as inverse functions of certain integrals. As for decompositions into squares, I recall that in Hardy and Wright's Introduction to the 'Theory of Numbers' there are three proofs of the four square theorem. One is 'elementary', one uses the integral quaternions/Hurwitz integers, and the third uses elliptic functions. This last one is (a newer version of ?) Jacobis proof.

These discoveries of connections between heterogeneous mathematical objects can be compared with the discovery of the connection between electricity and magnetism in physics or with the discovery of the similarity between the east coast of America and the west coast of Africa in geology. The emotional significance of such discoveries for teaching is difficult to overestimate. It is they who teach us to search and find such wonderful phenomena of harmony of the Universe. The de-geometrisation of mathematical education and the divorce from physics sever these ties. For example, not only students but also modern mathematicians on the whole do not know about the Jacobi fact mentioned here: an elliptic integral of first kind expresses the time of motion along an elliptic phase curve in the corresponding Hamiltonian system (see above). Rephrasing the famous words on the electron and atom, it can be said that a hypocycloid is as inexhaustible as an ideal in a polynomial ring. But teaching ideals to students who have never seen a hypocycloid is as ridiculous as teaching addition of fractions to children who have never cut (at least mentally) a cake or an apple into equal parts. No wonder that the children will prefer to add a numerator to a numerator and a denominator to a denominator. From my French friends I heard that the tendency towards super-abstract generalizations is their traditional national trait. I do not entirely disagree that this might be a question of a hereditary disease, but I would like to underline the fact that I borrowed the cake-and-apple example from Poincaré. The scheme of construction of a mathematical theory is exactly the same as that in any other natural science. First we consider some objects and make some observations in special cases. Then we try and find the limits of application of our observations, look for counter-examples which would prevent unjustified extension of our observations onto a too wide range of events (example: the number of partitions of consecutive odd numbers 1, 3, 5, 7, 9 into an odd number of natural summands gives the sequence 1, 2, 4, 8, 16, but then comes 29).

As a result we formulate the empirical discovery that we made (for example, the Fermat conjecture or Poincaré conjecture) as clearly as possible. After this there comes the difficult period of checking as to how reliable are the conclusions. At this point a special technique has been developed in mathematics. This technique, when applied to the real world, is sometimes useful, but can sometimes also lead to self-deception. This technique is called modelling. When constructing a model, the following idealisation is made: certain facts which are only known with a certain degree of probability or with a certain degree of accuracy, are considered to be absolutely correct and are accepted as 'axioms'. The sense of this absoluteness lies precisely in the fact that we allow ourselves to use these 'facts' according to the rules of formal logic, in the process declaring as 'theorems' all that we can derive from them. It is obvious that in any real-life activity it is impossible to wholly rely on such deductions. The reason is at least that the parameters of the studied phenomena are never known absolutely exactly and a small change in parameters (for example, the initial conditions of a process) can totally change the result. Say, for this reason a reliable long-term weather forecast is impossible and will remain impossible, no matter how much we develop computers and devices which record initial conditions. In exactly the same way a small change in axioms (of which we cannot be completely sure) is capable, generally speaking, of leading to completely different conclusions than those that are obtained from theorems which have been deduced from the accepted axioms. The longer and fancier is the chain of deductions

('proofs'), the less reliable is the final result. Complex models are rarely useful (unless for those writing their dissertations). The mathematical technique of modelling consists of ignoring this trouble and speaking about your deductive model in such a way as if it coincided with reality. The fact that this path, which is obviously incorrect from the point of view of natural science, often leads to useful results in physics is called 'the inconceivable effectiveness of mathematics in natural sciences' (or 'the Wigner principle'). Here we can add a remark by I.M. Gel'fand: there exists yet another phenomenon which is comparable in its inconceivability with the inconceivable effectiveness of mathematics in physics noted by Wigner - this is the equally inconceivable ineffectiveness of mathematics in biology. 'The subtle poison of mathematical education' (in Felix Klein's words) for a physicist consists precisely in that the absolutised model separates from the reality and is no longer compared with it. Here is a simple example: mathematics teaches us that the solution of the Malthus equation $dx/dt = x$ is uniquely defined by the initial conditions (that is that the corresponding integral curves in the (t,x) -plane do not intersect each other). This conclusion of the mathematical model bears little relevance to the reality. A computer experiment shows that all these integral curves have common points on the negative t -semi-axis. Indeed, say, curves with the initial conditions $x(0) = 0$ and $x(0) = 1$ practically intersect at $t = -10$ and at $t = -100$ you cannot fit in an atom between them. Properties of the space at such small distances are not described at all by Euclidean geometry. Application of the uniqueness theorem in this situation obviously exceeds the accuracy of the model. This has to be respected in practical application of the model, otherwise one might find oneself faced with serious troubles. Unfortunately, neither such examples, nor discussing the danger of fetishising theorems are to be met in modern mathematical textbooks, even in the better ones. I even got the impression that scholastic mathematicians (who have little knowledge of physics) believe in the principal difference of the axiomatic mathematics from modelling which is common in natural science and which always requires the subsequent control of deductions by an experiment. Not even mentioning the relative character of initial axioms, one cannot forget about the inevitability of logical mistakes in long arguments (say, in the form of a computer breakdown caused by cosmic rays or quantum oscillations). Every working mathematician knows that if one does not control oneself (best of all by examples), then after some ten pages half of all the signs in formulae will be wrong and twos will find their way from denominators into numerators. The technology of combatting such errors is the same external control by experiments or observations as in any experimental science and it should be taught from the very beginning to all juniors in schools. Attempts to create 'pure' deductive-axiomatic mathematics have led to the rejection of the scheme used in physics (observation - model - investigation of the model - conclusions - testing by observations) and its substitution by the scheme: definition - theorem - proof. It is impossible to understand an unmotivated definition.

A teacher of mathematics, who has not got to grips with at least some of the volumes of the course by Landau and Lifshitz, will then become a relict like the one nowadays who does not know the difference between an open and a closed set.

Translated by A.V. GORYUNOV

KAPITEL 1

Differenzierbare beschränkte finite Optimierung

§ 1 Motivation und Problemstellung

Wir betrachten zur Einführung nichtlineare, beschränkte Optimierungsprobleme, auch NLP (Nonlinear Program) genannt:

$$\left. \begin{array}{ll} \text{Minimiere} & f(x) \quad \text{über } x \in \mathbb{R}^n \\ \text{sodass} & g_i(x) \leq 0, \quad i = 1, \dots, k \\ \text{und} & h_j(x) = 0, \quad j = 1, \dots, m - k. \end{array} \right\} \quad (1.1)$$

mit $0 \leq k \leq m$.

Wir nehmen im folgenden an, dass f , g und h stetig differenzierbar auf \mathbb{R}^n sind.

$$S = \{x \in \mathbb{R}^n : g_i(x) \leq 0 \text{ für alle } i = 1, \dots, k, \quad h_j(x) = 0 \text{ für alle } j = 1, \dots, m - k\} \quad (1.2)$$

ist die **zulässige Menge** von (1.1).

Beispiel 1.1 (Gleichungsbeschränktes NLP1)

$$\begin{array}{ll} \min & f(x) = x_1 + x_2 \\ \text{unter} & c_1(x) = x_1^2 + x_2^2 - 2 = 0, \end{array} \quad (1.3)$$

d.h. $S = \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 - 2 = 0\}$.

Man berechnet $\nabla f(x) = (1, 1)^\top$, $\nabla c_1(x) = (2x_1, 2x_2)^\top$, die Gradientenvektoren sind orthogonal zu den Niveaulinien von c bzw. f . Visualisierung:

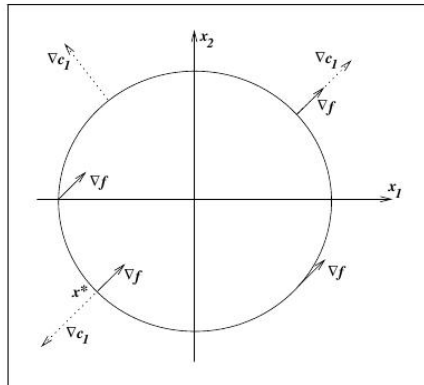


ABBILDUNG 1.1. aus Nocedal, Wright “Numerical Optimization”

Man sieht hier, dass in einem Punkt x^* immer sog. Abstiegsrichtungen (Richtungen, entlang derer die Funktionswerte von f lokal (!) kleiner werden) möglich sind, außer wenn ∇f und ∇c_1 parallel sind, d.h. wenn es ein $\lambda_1^* \in \mathbb{R}$ gibt mit $\nabla f(x^*) = \lambda_1^* \nabla c_1(x^*)$.

Dies ist also eine notwendige Bedingung für ein lokales Minimum von (1.3) !

Beispiel 1.2 (Ungleichungsbeschränktes NLP)

$$\begin{aligned} \min f(x) &= x_1 + x_2 \\ \text{unter } c_1(x) &= 2 - (x_1^2 + x_2^2) \geq 0, \end{aligned} \quad (1.4)$$

Man berechnet wieder $\nabla f(x) = (1, 1)^\top$, $\nabla c_1(x) = (-2x_1, -2x_2)^\top$, Visualisierung:

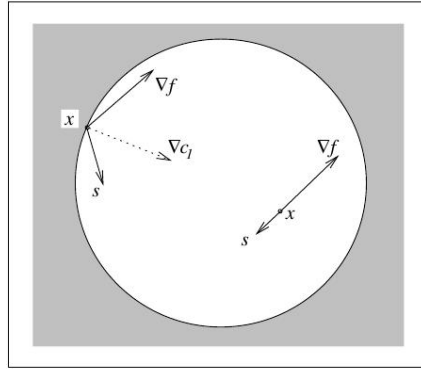


ABBILDUNG 1.2. aus Nocedal, Wright “Numerical Optimization”

Eine Abstiegsrichtung $s \in \mathbb{R}^2$ im Punkt $x^* \in \mathbb{R}^2$ muss die Bedingung $\langle \nabla f(x^*), s \rangle = \nabla f(x^*)^T s < 0$ erfüllen. Bzgl. der Zulässigkeit von s gilt in linearer Näherung (Taylor-Entwicklung 1. Ordnung):

$$0 \leq c_1(x^*) + \nabla c_1(x^*)^T s \quad (1.5)$$

Sei nun x^* eine optimale Lösung von (1.2) und s eine Abstiegsrichtung.

Fallunterscheidung:

- a) x^* liegt im Inneren der zulässigen Menge X , dann lässt sich immer eine Abstiegsrichtung s mit $\nabla f(x^*)^T s < 0$ finden, falls $\nabla f(x^*) \neq 0$, d.h. $\nabla f(x^*) = 0$ ist notwendige Bedingung für ein lokales Minimum.
- b) x^* liegt am Rand der zulässigen Menge X , dann gilt $c_1(x^*) = 0$ und mit (1.5) folgt

$$\nabla c_1(x^*)^T s \geq 0 \quad (1.6)$$

Zusammen mit $\nabla f(x^*)^T s < 0$ folgt also, dass $\nabla c_1(x^*)$ und $\nabla f(x^*)$ parallel sein und in die gleiche Richtung zeigen müssen, damit keine zulässige Abstiegsrichtung existiert (s. Abb. 1.3), d.h. $\exists \mu_1^* \in \mathbb{R}$ mit

$$\mu_1^* \nabla c_1(x^*) = \nabla f(x^*), \quad \mu_1 \geq 0 \quad (1.7)$$

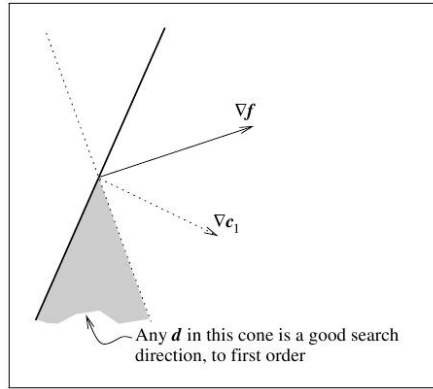


ABBILDUNG 1.3. Visualisierung Bedingung (1.7), aus Nocedal, Wright “Numerical Optimization”

§ 2 Einführung und allgemeine Konzepte der Optimierungstheorie

Die allgemeinste Formulierung von Optimierungsproblemen basiert auf dem abstrakten Konzept des topologischen Raums. Eine Topologie ist die Minimalstruktur, die nötig ist, um die Begriffe “offene Menge” und “Umgebung” charakterisieren zu können, die für die Definition eines “lokalen Extremums” benötigt werden.

Sei X ein topologischer Raum. $S \subset X$ Teilmenge und $f : X \rightarrow \mathbb{R}$ eine Abbildung.

Optimierungsproblem **(P)**:

$$\left. \begin{array}{ll} \text{Minimiere } f(x) \text{ unter der Nebenbedingung } x \in S \\ f : & \text{Zielfunktion} \\ S : & \text{zulässige Menge} \\ x \in S : & \text{zulässiger Punkt} \end{array} \right\} \quad (2.1)$$

a) Finite Optimierungsprobleme:

$$X = \mathbb{R}^n \text{ mit einer Norm } \|\cdot\|$$

b) Infinite Optimierungsprobleme:

$$X = B, \text{ } B \text{ Banachraum, z.B. } X = C[a, b] \text{ mit } \|\cdot\|_\infty$$

Definition 2.1

Ein Punkt $\bar{x} \in S \subset X$ heißt **lokale Minimalstelle** von **(P)**, falls es eine Umgebung $U \subset X$ mit $\bar{x} \in U$ gibt, so dass $f(\bar{x}) \leq f(x) \quad \forall x \in S \cap U$.

Maximierung einer Funktion f ist äquivalent zur Minimierung von $-f$.

Definition 2.2 (Value Function)

Der optimale Wert von **(P)** ist definiert als $v(P) := \inf\{f(x) : x \in S\}$, falls $S \neq \emptyset$ und $v(P) := \infty$, falls $S = \emptyset$.

Typen von Optimierungsproblemen:

- (1) Lineare Optimierungsprobleme
- (2) Nichtlineare Optimierungsprobleme
- (3) Konvexe Optimierungsprobleme
- (4) Nichtdifferenzierbare Optimierungsprobleme
- (5) Variationsprobleme / Optimalsteuerungsprobleme

Man unterscheidet zwischen kontinuierlichen (z.B. $S \subset \mathbb{R}$) und diskreten/ganzzahligen (z.B. $S \subset \mathbb{Z}$) Optimierungsproblemen je nach Struktur der Menge S . Diskrete Probleme erfordern aufgrund der Nichtdifferenzierbarkeit spezielle Theorie und numerische Lösungsverfahren, die i.d.R. kombinatorische Aspekte enthalten, und werden im Rahmen dieser Vorlesung nicht behandelt. Enthält ein Optimierungsproblem sowohl kontinuierliche als auch diskrete Variablen, spricht man von gemischt-ganzzahliger Optimierung.

Zentrale Aufgaben der Optimierungstheorie umfassen:

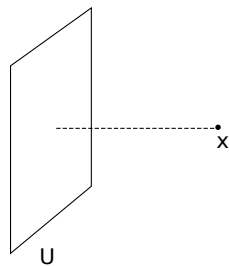
- Notwendige Optimalitätsbedingungen
- Hinreichende Optimalitätsbedingungen
- Sensitivitätsanalyse
- Effiziente numerische Verfahren

Einige geometrische Konzepte in der Optimierungstheorie

Die Optimierungstheorie ist i.w. geometrischer Natur und es ist i.a. hilfreich, sich einige zugrunde liegenden Konzepte zu veranschaulichen, bevor ein formaler und abstrakter begrifflicher Rahmen geschaffen wird.

Sei H ein Hilbertraum.

a) Projektionstheorem (Orthogonalprojektion im Hilbertraum)



Wichtig: Im Hilbertraum gilt Pythagoras mit dem Skalarprodukt als entscheidender Struktur und einem darauf basierenden Orthogonalitätsbegriff !

Das Projektionstheorem erlaubt die Charakterisierung des kürzesten Abstands (in der euklidischen Metrik) eines Punktes zu einem linearen Unterraum mit Hilfe der Orthogonalprojektion, die mit Hilfe des Skalarprodukts definiert wird.

Mit Hilfe des Dualraumkonzepts ist eine Verallgemeinerung von Hilberträumen auf Banachräume möglich. Ein wichtiger Satz ist in diesem Zusammenhang der Riesz'sche Darstellungssatz. In Kurzform lautet er

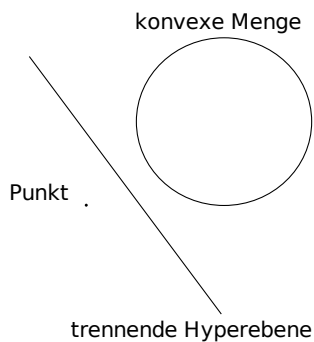
Riesz'scher Darstellungssatz:

Sei H ein reeller Hilbertraum mit dem Skalarprodukt $\langle \cdot, \cdot \rangle$ und H^* sein Dualraum, dann gibt es einen isometrischen Isomorphismus

$$\begin{aligned} H &\rightarrow H^* \\ v &\mapsto \langle v, \cdot \rangle \end{aligned}$$

In der Optimierungstheorie für Banachräume übernehmen die Elemente des Dualraums (lineare Funktionale auf H) die "Rolle" des Skalarprodukts in Hilberträumen.

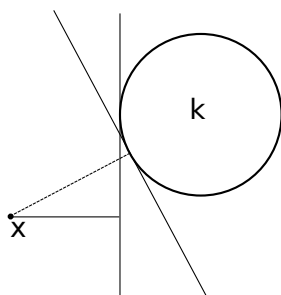
b) Trennungssatz ($\dim X < \infty$) bzw. Satz von Hahn-Banach ($\dim X = \infty$)



Der Trennungssatz für konvexe Mengen lautet:

Sei $K \subset X$ eine konvexe Menge und $y \notin \overline{K}$. Dann gibt es eine Hyperebene, die y und K trennt. Der Trennungssatz ist grundlegend für die Herleitung notwendiger Optimalitätsbedingungen beschränkter Optimierungsprobleme und in endlich- sowie unendlich-dimensionalen Räumen formulierbar.

c) Dualität



Duale Formulierung: „Kleinsten Abstand von x zur Menge K ist gleich dem größten Abstand von x zu allen möglichen x und K trennenden Hyperebenen.“

Für "primale" Optimierungsprobleme lassen sich in vielen Fällen "duale" Probleme formulieren, die eine komplementäre Struktur haben und deren Lösungen im gewissen Sinne denen der primalen Probleme entsprechen. Primal-duale Formulierungen sind von großer Bedeutung für die Optimierungstheorie und numerische Lösungsverfahren.

Fragestellungen bei Optimierungsaufgaben:

- Wann existieren Optimallösungen ?
- Wie überprüft man Optimalität ? (\leadsto Optimalitätsbedingungen)
- Wie kann man Optimallösungen näherungsweise numerisch bestimmen ?
- Wie konstruiert man effiziente Algorithmen?

Literatur:

1. Optimierung. Florian Jarre und Josef Stoer, Springer, Berlin 2003
2. Nichtlineare Optimierung: Eine Einführung in Theorie, Verfahren und Anwendungen. Walter Alt, Vieweg und Teubner, 2002
3. **Theorie und Numerik restringierter Optimierungsaufgaben.** Christian Kanzow und Carl Geiger, Springer, Berlin 2002
4. Numerische Verfahren zur Lösung unrestringierter Optimierungsaufgaben. Christian Kanzow und Carl Geiger, Springer, Berlin 1999
5. Linear and Nonlinear Programming (International Series in Operations Research and Management Science). David G. Luenberger und Yinyu Ye, Springer, Berlin 2008
6. Optimization by Vector Space Methods. David G. Luenberger, Wiley and Sons 1998
7. **Numerical Optimization** (Springer Series in Operations Research and Financial Engineering). Jorge Nocedal und Stephen J. Wright, Springer, Berlin 2006
8. **Nichtlineare Optimierung.** S. und M. Ulbrich, Birkhäuser 2012
9. Numerik der Optimierung. Christian Großmann und Johannes Terno, Teubner, 1997
10. Practical Methods of Optimization. Roger Fletcher, Wiley VCH 2000
11. Convex Optimization. Stephen Boyd und Lieven Vandenberghe, Cambridge University Press 2004
12. Nonlinear Programming. Lorenz T. Biegler, SIAM 2010
13. Optimization. Elijah Polak, Springer 1997
14. **Perturbation Analysis of Optimization Problems. J.Frederic Bonnans und Alexander Shapiro, Springer 2000**
15. Optimal Control of ODEs and DAEs. M. Gerds, De Gruyter 2012
16. Practical Methods for Optimal Control Using Nonlinear Programming. J.T. Betts, SIAM 2001.

In **Fettdruck** die für diese Vorlesung empfohlenen Titel, 14. ist anspruchsvoll !

§ 2.1 Existenz von Optimallösungen

Der Beweis der Existenz von Lösungen eines Optimierungsproblems läuft i.d.R. prinzipiell über den Satz von Weierstrass.

Satz 2.3 (Weierstrass)

Sei $S \subset X$, $S \neq \emptyset$ kompakt und $f : S \rightarrow \mathbb{R}$ stetig. Dann besitzt das Optimierungsproblem (2.1) mindestens eine Lösung.

Beweis: siehe Analysis II.

□

In endlich-dimensionalen Räumen sichert der Satz von Heine-Borel die Kompaktheit von beschränkten und abgeschlossenen Mengen. Diese Äquivalenz ist jedoch im unendlich-dimensionalen Fall nicht mehr gültig.

In diesem Fall kommt häufig das Konzept der schwachen Kompaktheit zum Tragen. Eine Menge heißt schwach kompakt, wenn Sie kompakt bezüglich der schwachen Topologie ist. Offene Mengen werden in der schwachen Topologie wie folgt konstruiert.

Beispiel 2.4 (Kompaktheit in ∞ -dim. Räumen)

Wähle $X = (C([a, b]), \|\cdot\|_\infty)$ mit $[a, b] := [0, 1]$. Definiere

$$f_n : [0, 1] \rightarrow \mathbb{R}, \quad x \mapsto \begin{cases} 1 - nx & , x < \frac{1}{n} \\ 0 & , \text{sonst.} \end{cases}$$

Dann gilt $f_n \in X$ für alle $n \in \mathbb{N}$ und weiter

$$\|f_n\|_\infty = \sup_{x \in [0, 1]} |f_n(x)| = 1 \quad \forall n \in \mathbb{N} \quad \Leftrightarrow \quad (f_n) \text{ beschränkt.}$$

Die Folge konvergiert nicht gleichmäßig bzgl. der Supremumsnorm, denn für $m \geq n$ gilt

$$\begin{aligned} \|f_n - f_m\|_\infty &= \max \left\{ \max_{x \in [0, \frac{1}{m}]} \{|1 - nx - 1 + mx|\}, \max_{x \in [\frac{1}{m}, \frac{1}{n}]} \{|1 - nx|\}, 0 \right\} \\ &= \max \left\{ (m - n) \cdot \frac{1}{m}, 1 - n \frac{1}{m}, 0 \right\} = 1 - \frac{n}{m} \quad \forall m \geq n \in \mathbb{N} \end{aligned}$$

und daher insbesondere

$$\|f_n - f_{2n}\| = \frac{1}{2} \quad \forall n \in \mathbb{N}.$$

Der punktweise Limes $f : [0, 1] \rightarrow \mathbb{R}$ mit

$$\lim_{n \rightarrow \infty} f_n(x) =: f(x) = \begin{cases} 1 & , x = 0 \\ 0 & , \text{sonst} \end{cases}$$

ist unstetig. Daraus folgt, dass es in X keine konvergente Teilfolge von $(f_n)_{n \in \mathbb{N}}$ gibt und somit ist die Einheitsphäre

$$S_1 := \{g \in C([0, 1]) : \|g\|_\infty = 1\}$$

nicht kompakt.

Definition 2.5 (Schwache Topologie)

Sei X ein normierter Raum und $S \subset X$. S heißt offen (bzgl. der schwachen Topologie), falls es eine offene Menge $O \subset \mathbb{R}$ und ein normstetiges lineares Funktional Φ auf X (d.h. eine lineare, stetige Abbildung $\Phi : X \rightarrow \mathbb{R}$) gibt, so dass $S = \Phi^{-1}O$ oder S endlicher Durchschnitt oder beliebige Vereinigung derartiger Urbilder ist.

Definition 2.6 (Schwache Konvergenz)

Eine Folge $\{x^k\}_{k \in \mathbb{N}}$ in einem normierten Raum X heißt schwach konvergent (d.h. konvergent in der schwachen Topologie) gegen $\bar{x} \in X$, wenn gilt

$$\lim_{k \rightarrow \infty} \Phi(x^k) = \Phi(\bar{x}) \quad \forall \Phi \in X^* \text{ (Dualraum)}$$

Beispiel 2.7 (Schwache / starke Konvergenz)

Sei der Hilbertraum $X := L^2(0, 2\pi)$ gegeben und

$$f_n(x) := \frac{1}{\sqrt{\pi}} \sin(nx).$$

Dann bildet $\{f_n : n \in \mathbb{N}\}$ eine Orthogonalbasis (Fourierbasis) von X . Sei $f \in X$ beliebig. Dann liefert die Projektion

$$\langle f, f_n \rangle = \frac{1}{\sqrt{\pi}} \int_0^{2\pi} f(x) \sin(nx) dx$$

den n -ten Koeffizient in der Basisentwicklung von f nach f_n (also den n -ten Fourierkoeffizient). Aus der Parseval'schen Gleichung

$$\sum_{n=1}^{\infty} |\langle f, f_n \rangle|^2 = \|f\|^2$$

folgt

$$\langle f, f_n \rangle \rightarrow 0 \quad , n \rightarrow \infty.$$

Wegen des Riesz'schen Darstellungssatzes konvergiert die Folge (f_n) also schwach gegen $0 = \langle f, 0 \rangle$, aber ferner gilt

$$\|f_n - 0\|^2 = \frac{1}{\pi} \int_0^{2\pi} \sin^2(nx) dx = 1,$$

also konvergiert die Folge (f_n) **nicht** stark gegen 0.

Da die schwache Stetigkeit (bzgl. der schwachen Topologie) jedoch i.a. eine zu starke Voraussetzung ist, um für Existenzbeweise nützlich zu sein, definiert man noch:

Definition 2.8 (Unterhalbstetigkeit)

Sei X ein normierter Raum, $\bar{x} \in S \subset X$, $f : S \rightarrow \mathbb{R}$. Die Funktion f heißt (schwach) unterhalbstetig in $\bar{x} \in S$, falls für jede (schwach) gegen \bar{x} konvergente Folge $\{x^k\}$ gilt

$$f(\bar{x}) \leq \liminf_{k \rightarrow \infty} f(x^k)$$

Die Bedeutung dieser Begriffe zeigt der folgende Satz.

Satz 2.9

Sei X ein Banachraum kompakt und $f : X \rightarrow \mathbb{R}$ **konvex** und stetig. Dann ist f schwach unterhalbstetig, d.h. für eine Folge $\{x^k\}_{k \in \mathbb{N}}$, die in schwach gegen $\bar{x} \in X$ konvergiert, gilt $f(\bar{x}) \leq \liminf_{k \rightarrow \infty} f(x^k)$.

Damit kann man eine geeignete Modifikation des Satzes von Weierstrass formulieren, der eine Existenzaussage von Optimallösungen unter schwächeren Bedingungen liefert.

Korollar 2.10 (zum Satz von Weierstrass)

Es existiere ein $\bar{x} \in S$, so dass die Niveaumenge

$$W(\bar{x}) := \{x \in S : f(x) \leq f(\bar{x})\}$$

(schwach) kompakt und f (schwach) unterhalbstetig auf $W(\bar{x})$ ist. Dann besitzt das Optimierungsproblem (2.1) mindestens eine Lösung.

In Banachräumen, in denen der Satz von Heine-Borel nicht gilt, kann man mit Hilfe der schwachen Kompaktheit und Korollar 2.10 Existenzaussagen machen, sobald man hinreichende Bedingungen zur Verfügung hat, welche die schwache Kompaktheit von Mengen in diesen Räumen sicher stellen. Eine wichtige Eigenschaft im Zusammenhang mit (starker) Kompaktheit ist, dass konvexe, abgeschlossene und beschränkte Mengen in reflexiven Banachräumen kompakt sind.

Die in diesem Paragraphen kurz eingeführten Konzepte spielen für die Theorie der Unendlichen Optimierung eine wichtige Rolle. Sie kommen insbesondere bei der Optimalsteuerung partieller Differentialgleichungen zum Einsatz.

§ 2.2 Grundbegriffe der Optimierung im \mathbb{R}^n

Allgemeines beschränktes Optimierungsproblem

$$\left. \begin{array}{ll} \text{Minimiere} & f(x) \quad \text{über } x \in \Omega \quad (\text{Zielfunktion}) \\ \text{sodass} & h_i(x) = 0, \quad i \in \mathcal{E} \quad (\text{Gleichungsnebenbedingungen}) \\ \text{und} & g_i(x) \leq 0, \quad i \in \mathcal{I} \quad (\text{Ungleichungsnebenbedingungen}) \end{array} \right\} \quad (2.2)$$

Ω heißt **Grundmenge** und x die **Variable(n)** der Aufgabe. Wir behandeln im ersten Teil der Vorlesung i.w. Optimierungsaufgaben der Form

- $\Omega \subset \mathbb{R}^n$, Ω konvex (meist $\Omega = \mathbb{R}^n$)
- $f, g_i, h_i : \mathbb{R}^n \rightarrow \mathbb{R}$ hinreichend glatte Funktionen
- \mathcal{E} und \mathcal{I} endliche (ggf. leere) Indextmengen.

Definition 2.11

Die Menge

$$X = \{x \in \Omega : h_i(x) = 0 \text{ für alle } i \in \mathcal{E}, \quad g_i(x) \leq 0 \text{ für alle } i \in \mathcal{I}\}$$

heißt **zulässige Menge**, $x \in X$ **zulässiger Punkt**.

- (a) Ein $x^* \in X$ heißt **globales Optimum**, wenn gilt:

$$f(x^*) \leq f(x) \text{ für alle } x \in X.$$

- (b) Ein globales Optimum heißt **strikt**, wenn gilt:

$$f(x^*) < f(x) \text{ für alle } x \in X, \quad x \neq x^*.$$

- (c) Ein $x^* \in X$ heißt **lokales Optimum**, wenn es eine Umgebung $U(x^*)$ gibt, sodass gilt:

$$f(x^*) \leq f(x) \text{ für alle } x \in X \cap U(x^*).$$

- (d) Ein lokales Optimum heißt **strikt**, wenn gilt:

$$f(x^*) < f(x) \text{ für alle } x \in X \cap U(x^*), \quad x \neq x^*.$$

Ein Optimum nennt man auch **Minimum**, **Minimalstelle** oder **Lösung** von (2.2).

§ 2.3 Klassifikation von Optimierungsproblemen

Definition 2.12 (Klassifikation von Optimierungsproblemen)

- (a) Die Optimierungsaufgabe (2.2) heißt **frei** oder **unrestringiert**, wenn $\mathcal{E} = \mathcal{I} = \emptyset$ ist, andernfalls **gleichungs-** und/oder **ungleichungs-restringiert** oder **-beschränkt**.

- (b) Ungleichungsbeschränkungen der Art

$$l_i \leq x_i \leq u_i, \quad i = 1, \dots, n$$

mit $l_i \in \mathbb{R} \cup \{-\infty\}$ und $u_i \in \mathbb{R} \cup \{\infty\}$ heißen **Box-Beschränkungen**.

- (c) Sind f , g und h (affin-)lineare Funktionen von x , so sprechen wir von **linearer Optimierung**. Ein lineares Optimierungsproblem heißt auch **lineares Programm (LP)**, also z.B.

$$\min \quad c^\top x \quad \text{sodass} \quad Ax = b \quad \text{und} \quad x \geq 0.$$

- (d) Sind allgemeiner f und alle g_i konvexe Funktionen und sind alle h_i wieder (affin-)linear, so sprechen wir von **konvexer Optimierung**.

- (e) Ist f ein quadratisches Polynom und sind g und h (affin-)linear, so sprechen wir von **quadratischer Optimierung**. Ein quadratisches Optimierungsproblem heißt auch **quadratisches Programm (QP)**.

- (f) Im allgemeinen Fall spricht man von **nichtlinearer Optimierung** und von einem **nichtlinearen Programm (NLP)**.

§ 2.4 Beispiele für Optimierungsprobleme

Häufig ist die “Modellierung”, d.h. die mathematische Formulierung der erste Schritt zur Lösung einer Aufgabenstellung. Beispiele für praxismotivierte Optimierungsprobleme sind z.B.

Beispiel 2.13 (Ökonomische Kosten)

Ein Unternehmen will eine bestimmte Menge M einer Ware einkaufen und holt dazu Angebote von n Lieferfirmen ein, von denen keine die gewünschte Gesamtmenge alleine liefern kann. Anbieter i liefert maximal m_i , wobei der Preis $f_i(x_i)$ von der Bestellmenge x_i abhängt. [f_i wird i.d.R. nichtlinear sein.] Die minimalen Beschaffungskosten ergeben sich durch Lösung der folgenden Optimierungsaufgabe:

$$\left. \begin{array}{ll} \min & f(x) := \sum_{i=1}^n f_i(x_i), \quad x \in \mathbb{R}^n \\ \text{sodass} & \sum_{i=1}^n x_i = M \\ \text{und} & 0 \leq x_i \leq m_i, \quad i = 1, \dots, n. \end{array} \right\} \quad (2.3)$$

Beispiel 2.14 (Parameteridentifizierungsproblem)

Eine an einer Feder befestigte Masse bewegt sich unter dem Einfluss von Dämpfung gemäß der Differentialgleichung

$$m\ddot{y} + r\dot{y} + ky = 0.$$

In einem Experiment mit Anfangsanregung

$$y(0) = y_0, \quad \dot{y}(0) = 0$$

soll zu verschiedenen Zeitpunkten t_i , $i = 1, \dots, N$ die Auslenkung y_i gemessen werden. Daraus sollen bei bekannter Masse m die unbekannten Dämpfungs- und Federkonstanten (Parameter) r und k bestimmt werden.

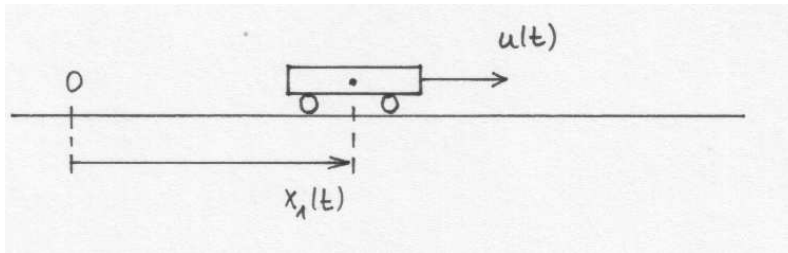
Setze $\theta := (r, k)^\top$. Die Lösung der Differentialgleichung zum Zeitpunkt t mit obigen Anfangsbedingungen wird mit $y(t; \theta)$ bezeichnet, um die Abhängigkeit von den zu bestimmenden Parametern θ zu verdeutlichen. Wir formulieren folgende Aufgabe:

$$\min_{\theta \in \mathbb{R}^2} f(\theta) := \frac{1}{2} \sum_{i=1}^N |y(t_i; \theta) - y_i|^2, \quad (2.4)$$

damit die Messwerte y_i möglichst gut die Vorhersage der Position durch das Modell $y(t_i; \theta)$ wiedergeben. Motivation für diese „Methode der kleinsten Quadrate“: Sind die Messfehler in den Daten y_i unabhängig und identisch normalverteilt, so erhält man aus (2.4) den wahrscheinlichsten Parameterwert θ .

Parameterschätzprobleme sind sog. semi-infinite Optimierungsprobleme, der Raum der Optimierungsvariablen (Parameterraum) ist endlich-dimensional, aber es gibt unendlich viele Beschränkungen (Differentialgleichung).

Beispiel 2.15 (Optimalsteuerungsproblem)



Problemstellung: Der Wagen soll möglichst schnell zum Nullpunkt gesteuert werden und dort stehenbleiben. Es seien

$x_1(t)$: Position des Wagens (Masse = 1) zur Zeit t

$x_2(t)$: Geschwindigkeit des Wagens

$u(t)$: Beschleunigung des Wagens, **Steuerfunktion**

$x(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}$: **Zustand** des Systems.

Die Dynamik des Systems ist nach dem zweiten Newton'schen Axiom (Kraft = Masse \times Beschleunigung) gegeben durch

$$\left. \begin{array}{ll} \dot{x}_1(t) = x_2(t) \\ \dot{x}_2(t) = u(t) \\ x(0) = x_0 \in \mathbb{R}^2 & \text{Anfangsbedingung} \\ x(T) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} & \text{Endbedingung} \\ u(t) \in [-1, 1] & \text{Steuerbeschränkung} \end{array} \right\} \quad (2.5)$$

Ziel: Bestimme eine stückweise stetige optimale Steuerung $u : [0, T] \rightarrow \mathbb{R}$ so, dass die Endzeit T minimal wird unter den Beschränkungen (2.5).

Optimalsteuerungsprobleme sind unendlich-dimensionale (infinite) Optimierungsprobleme, da die Optimierungsvariablen (Steuer-)Funktionen sind, also Elemente eines unendlich-dimensionalen Raums. Man kann sie mit Hilfe sog. direkter numerischer Methoden behandeln, indem man die Steuerfunktionen parametrisiert (z.B. durch Polynome approximiert) und dadurch auf einen endlich-dimensionalen Raum projiziert.

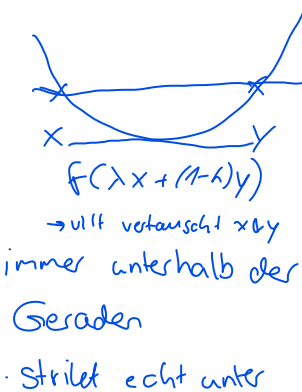
§ 2.5 Konvexität und Trennungssatz im \mathbb{R}^n

Definition 2.16

↳ Theorie für Nebenbedingungen



konvexe Fkt:



(a) Eine Menge $K \subset \mathbb{R}^n$ heißt **konvex**, wenn gilt:

$$\lambda f(x) + (1-\lambda)f(y) \quad x, y \in K \Rightarrow \alpha x + (1-\alpha)y \in K, \quad 0 \leq \alpha \leq 1.$$

(b) Die **Konvexe Hülle** von $M \subset \mathbb{R}^n$ ist definiert als

$$\text{co}(M) := \bigcap_{K \subset \mathbb{R}^n} \{K \mid M \subset K, K \text{ konvex}\}.$$

Das ist die kleinste konvexe Teilmenge des \mathbb{R}^n , die M enthält. (Wird in der Literatur auch häufig mit $\text{conv}(M)$ bezeichnet.)

(c) Eine Menge $K \subset \mathbb{R}^n$ heißt **Kegel** (mit Spitze in 0), wenn $\alpha x \in K, \forall x \in K, \alpha \geq 0$. *

(d) Sei $K \subset \mathbb{R}^n$ und $x \in K$. Der Kegel $K(x) := \{\alpha(y - x) \mid y \in K, \alpha > 0\} = \bigcup_{\alpha > 0} \alpha(K - x)$ heißt **konische Hülle (Kegelhülle)** von K bzgl. x . (Wird in der Literatur auch häufig mit $\text{cone}(K - x)$ bezeichnet.)

Definition 2.17 (Konvexkombination)

(a) $x \in \mathbb{R}^n$ heißt **Konvexkombination** von $x_1, \dots, x_m \in \mathbb{R}^n$, falls $x = \sum_{i=1}^m \lambda_i x_i$ gilt mit Skalaren $\lambda_i \geq 0$ und $\sum_{i=1}^m \lambda_i = 1$. Die Konvexkombination heißt **echt**, wenn alle $\lambda_i > 0$ sind.

(b) Ist $X \subset \mathbb{R}^n$ eine (nicht notwendig endliche) Menge, so heißt x eine Konvexkombination von X , wenn es endlich viele Vektoren $x_1, \dots, x_m \in X$ gibt, sodass x eine Konvexkombination dieser Vektoren ist. \diamond

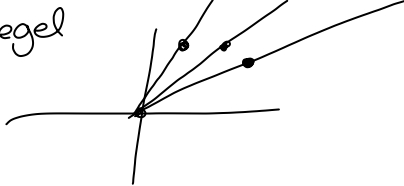
Lemma 2.18

Eine Menge $K \subset \mathbb{R}^n$ ist genau dann konvex, wenn sie alle Konvexkombinationen ihrer Elemente enthält.

Beweis: „ \Rightarrow “: Sei K konvex. Für $m \in \mathbb{N}$ und $x_1, \dots, x_m \in X$ sowie $\lambda_1, \dots, \lambda_m \geq 0$ mit $\sum_{i=1}^m \lambda_i = 1$ ist zu zeigen: $x = \sum_{i=1}^m \lambda_i x_i \in K$.

Induktion nach m : Für $m = 1, 2$ ist die Behauptung erfüllt. Es sei bereits gezeigt, dass K alle Konvexkombinationen von höchstens m Elementen enthält.

* z.B. μ konvexe Hülle minimal es auffüllen um Menge K konvex zu machen



Schluss auf $m+1$: Sei $\sum_{i=1}^{m+1} \lambda_i = 1, \lambda_i \geq 0, i = 1, \dots, m+1$. O.B.d.A. gilt $\lambda_{m+1} < 1$. [Ansonsten ist $x = x_{m+1}$ und nichts zu zeigen.] Setze $\alpha_i := \frac{\lambda_i}{1-\lambda_{m+1}}$ für $i = 1, \dots, m$. Dann ist $\alpha_i \geq 0$ und $\sum_{i=1}^m \alpha_i = 1$. Der Vektor $y = \sum_{i=1}^m \alpha_i x_i$ gehört nach Induktionsvoraussetzung zu K , also auch $x = (1 - \lambda_{m+1})y + \lambda_{m+1}x_{m+1}$ wegen der Konvexität von K .

„ \Leftarrow “: Seien $x_1, x_2 \in K, \lambda \in [0, 1]$. Nach Voraussetzung enthält K alle Konvexkombinationen $\lambda x_1 + (1 - \lambda)x_2$, d.h. K ist konvex. \square

Lemma 2.19 (Konvexkombination)

Die konvexe Hülle $\text{co}(M)$ einer nichtleeren Teilmenge $M \subset \mathbb{R}^n$ ist gegeben durch

$$\text{co}(M) = \left\{ \sum_{i=1}^k \alpha_i x_i : x_i \in M, \alpha_i \geq 0, \sum_{i=1}^k \alpha_i = 1, k \in \mathbb{N} \right\}$$

Beweis: Sei Y die Menge auf der rechten Seite.

„ $\text{co}(M) \subset Y$ “: Wir zeigen, dass Y konvex ist. Dann kommt diese Menge im Durchschnitt (Def. 2.16 b) vor, also gilt schon $\text{co}(M) \subset Y$.

Seien $x, y \in Y$, also gibt es Zahlen $m, l \in \mathbb{N}$ und $\lambda_1, \dots, \lambda_m \geq 0$ sowie $\mu_1, \dots, \mu_l \geq 0$ mit $\sum_{i=1}^m \lambda_i = 1$ und $\sum_{j=1}^l \mu_j = 1$, sodass $x = \sum_{i=1}^m \lambda_i x_i$ und $y = \sum_{j=1}^l \mu_j y_j$ gelten. Sei $\lambda \in [0, 1]$. Dann gilt

$$\lambda x + (1 - \lambda)y = \lambda \sum_{i=1}^m \lambda_i x_i + (1 - \lambda) \sum_{j=1}^l \mu_j y_j,$$

d.h., $\lambda x + (1 - \lambda)y$ ist Linearkombination der $\{x_i\}_{i=1}^m \cup \{y_j\}_{j=1}^l$. Die Koeffizienten sind ≥ 0 und ergeben in der Summe 1. Damit ist $\lambda x + (1 - \lambda)y \in Y$ und Y konvex.

„ $\text{co}(M) \supset Y$ “: $\text{co}(M)$ ist per Definition konvex, enthält daher nach Lemma 2.18 alle Konvexkombinationen ihrer eigenen Elemente. Also enthält $\text{co}(M)$ insbesondere alle Elemente aus Y , die ja Konvexkombinationen von Elementen aus M sind. \square

Der **Satz von Caratheodory** zeigt, dass für die Darstellung jedes Elements von $\text{co}(M)$ höchstens $(n+1)$ Elementen aus $M \subset \mathbb{R}^n$ benötigt werden (siehe Übungsaufgabe).

Lemma 2.20

Seien $C \subset \mathbb{R}^n$ konvex, $x_1 \in \overline{C}$ und $x_2 \in \text{int}(C)$. Dann gilt $\lambda x_1 + (1 - \lambda)x_2 \in \text{int}(C)$ für alle $\lambda \in [0, 1)$, d.h., die gesamte Verbindungsstrecke (evtl. mit Ausnahme von x_1 selbst) gehört zum Inneren von C .

Beweis: Es existiert $\varepsilon > 0$ mit $U_\varepsilon(x_2) \subset C$. Sei $\lambda \in (0, 1)$ [Fall $\lambda = 0$ trivial] und $y = \lambda x_1 + (1 - \lambda)x_2$. Wir zeigen: $U_r(y) \subset C$ für $r := (1 - \lambda)\varepsilon$. Damit ist gezeigt: $y \in \text{int}(C)$.

Sei $z \in U_r(y)$, also $\|y - z\| < (1 - \lambda)\varepsilon$. Wegen $x_1 \in \overline{C}$ [Häufungspunkt einer Folge aus C !] ist $U_\delta(x_1) \cap C \neq \emptyset$ für alle $\delta > 0$. Setze

$$\delta := \frac{(1 - \lambda)\varepsilon - \|y - z\|}{\lambda} > 0.$$

Es existiert also $z_1 \in U_\delta(x_1) \cap C$, d.h.,

$$z_1 \in C \quad \text{und} \quad \|z_1 - x_1\| < \frac{(1 - \lambda)\varepsilon - \|y - z\|}{\lambda}.$$

Ziel: Schreibe z als Konvexkombination von z_1 und einem weiteren $z_2 \in C$, nämlich einem $z_2 \in U_\varepsilon(x_2)$. Setze dazu

$$z_2 := \frac{1}{1-\lambda}z - \frac{\lambda}{1-\lambda}z_1 = \frac{z - \lambda z_1}{1-\lambda}.$$

Dann gilt $z = \lambda z_1 + (1-\lambda)z_2$ und

$$\begin{aligned} z_2 - x_2 &= \frac{z - \lambda z_1}{1-\lambda} - x_2 = \frac{z - \lambda z_1 - (y - \lambda x_1)}{1-\lambda} \\ \Rightarrow \|z_2 - x_2\| &\leq \frac{1}{1-\lambda}(\|z - y\| + \lambda \|x_1 - z_1\|) \\ &< \frac{1}{1-\lambda}(\|z - y\| + (1-\lambda)\varepsilon - \|z - y\|) \\ &= \varepsilon, \end{aligned}$$

also $z_2 \in U_\varepsilon(x_2) \subset C$. Aus $z = \lambda z_1 + (1-\lambda)z_2$ folgt $z \in C$, was zu zeigen war. \square

Definition 2.21 (Konvexe Funktion)

Es sei $C \subset \mathbb{R}^n$ konvex, $C \neq \emptyset$. Eine Funktion $f : C \rightarrow \mathbb{R}$ heißt

(a) **konvex** auf C , falls

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$$

gilt für alle $x, y \in C$ und $\lambda \in [0, 1]$.

(b) **strikt konvex** auf C , falls

$$f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda)f(y)$$

gilt für alle $x, y \in C$ mit $x \neq y$ und $\lambda \in (0, 1)$.

(c) **gleichmäßig konvex** auf C , wenn es eine Konstante $\mu > 0$ gibt mit

funktion $\underbrace{f(\lambda x + (1-\lambda)y)} + \underbrace{\mu \lambda (1-\lambda) \|x - y\|^2} \leq \underbrace{\lambda f(x) + (1-\lambda)f(y)}_{\text{Gerade}}$
gilt für alle $x, y \in C$ und $\lambda \in [0, 1]$. *min Abstand von fkt zur Gerade*

(d) [strikt, gleichmäßig] **konkav** auf C , wenn $-f$ [strikt, gleichmäßig] konvex auf C ist. \diamond

Beachte: f gleichmäßig konvex $\Rightarrow f$ strikt konvex $\Rightarrow f$ konvex

Beispiel 2.22 (Beispiele konvexer Funktionen)

(a) Affin-lineare Funktionen $f(x) = a^\top x + \beta$ sind konvex und konkav auf \mathbb{R}^n .

(b) Quadratische Funktionen $f(x) = \frac{1}{2}x^\top Qx + c^\top x + \gamma$ mit $Q = Q^\top \in \mathbb{R}^{n \times n}$:
 f konvex $\Leftrightarrow Q$ ist positiv semidefinit,

f gleichmäßig konvex $\Leftrightarrow Q$ ist positiv definit.

(c) $f(x) = \|x - y\|$ mit $y \in \mathbb{R}^n$ ist konvex auf \mathbb{R}^n , aber nicht strikt konvex.

(d) $f(x) = \|x - y\|^2$ mit $y \in \mathbb{R}^n$ ist gleichmäßig konvex auf \mathbb{R}^n .

(e) $f(x) = \|x - y\|^4$ mit $y \in \mathbb{R}^n$ ist nur strikt konvex.

(f) $f(x) = \exp(x)$ ist strikt konvex auf \mathbb{R} , aber nicht gleichmäßig konvex.

(g) $f(x) = \log(x)$ ist strikt konkav auf $(0, \infty)$. \diamond

Satz 2.23 (Operationen auf konvexen Funktionen)

Sei $C \subset \mathbb{R}^n$ konvex.

- (a) $f_i : C \rightarrow \mathbb{R}$ konvex und $\alpha_i \geq 0$ für $i = 1, \dots, m$. Dann ist

$$f(x) := \sum_{i=1}^m \alpha_i f_i(x)$$

konvex auf C . Ist mindestens ein f_i strikt konvex und das zugehörige $\alpha_i > 0$, so ist f strikt konvex auf C .

- (b) Ist $g : C \rightarrow \mathbb{R}^m$ affin und $f : \underbrace{g(C)}_{\text{konvex}} \rightarrow \mathbb{R}$ konvex, so ist $(f \circ g)$ konvex auf C .
- (c) Ist $f : C \rightarrow \mathbb{R}$ konvex, so ist auch $g(x) := \max\{0, f(x)\}$ konvex auf C .
- (d) Ist $f : C \rightarrow \mathbb{R}$ konvex und $f(x) \geq 0$, so ist $g(x) := f(x)^2$ konvex auf C .

Beweis: Übungsaufgabe. □

Satz 2.24 (Charakterisierung konvexer Funktionen)

Sei $C \subset \mathbb{R}^n$ offen und konvex sowie $f : C \rightarrow \mathbb{R}$ [stetig] differenzierbar. Dann gelten:

- (a) f ist konvex auf $C \Leftrightarrow$ Für alle $x, y \in C$ gilt:

$$f(x) - f(y) \geq \nabla f(y)^\top (x - y). \quad (2.6)$$

- (b) f ist strikt konvex auf $C \Leftrightarrow$ Für alle $x, y \in C$ mit $x \neq y$ gilt:

$$f(x) - f(y) > \nabla f(y)^\top (x - y). \quad (2.7)$$

- (c) f ist gleichmäßig konvex auf $C \Leftrightarrow$ Es existiert $\mu > 0$, sodass für alle $x, y \in C$ gilt:

$$f(x) - f(y) \geq \nabla f(y)^\top (x - y) + \mu \|x - y\|^2. \quad (2.8)$$

Beachte: Nach Teil (a) ist eine C^1 -Funktion $f : C \rightarrow \mathbb{R}$ genau dann konvex, wenn in jedem Punkt $y \in C$ die Tangente

$$L(x) := f(y) + \nabla f(y)^\top (x - y)$$

nicht oberhalb des Graphen von f verläuft.

Beweis: Zunächst „ \Leftarrow “ im Fall (c): Es gelte (2.8). Seien $x, y \in C$ und $\lambda \in (0, 1)$. Setze $z := \lambda x + (1 - \lambda)y$. Es folgt durch Addition von

$$f(x) - f(z) \geq \nabla f(z)^\top (x - z) + \mu \|x - z\|^2 \quad \Big| \cdot \lambda$$

und

$$f(y) - f(z) \geq \nabla f(z)^\top (y - z) + \mu \|y - z\|^2 \quad \Big| \cdot (1 - \lambda)$$

$$\begin{aligned} \Rightarrow \quad & \lambda f(x) + (1 - \lambda)f(y) - f(z) \geq \underbrace{\lambda \nabla f(z)^\top x + (1 - \lambda) \nabla f(z)^\top y - \nabla f(z)^\top z}_{=0} \\ & + \underbrace{\mu \lambda \|x - z\|^2 + \mu (1 - \lambda) \|y - z\|^2}_{= \mu \lambda (1 - \lambda) \|x - y\|^2} \quad (\text{nachrechnen}) \end{aligned}$$

$$\Rightarrow \quad \lambda f(x) + (1 - \lambda)f(y) - f(\lambda x + (1 - \lambda)y) \geq \mu \lambda (1 - \lambda) \|x - y\|^2,$$

d.h., f ist gleichmäßig konvex.

Analog zeigt man „ \Leftarrow “ im Fall (a) und (b).

„ \Rightarrow “ im Fall (c): Sei f gleichmäßig konvex und $x, y \in C$, $\lambda \in (0, 1)$. Dann gilt

$$\begin{aligned} f(y + \lambda(x - y)) &= f(\lambda x + (1 - \lambda)y) \\ &\leq \lambda f(x) + (1 - \lambda)f(y) - \mu \lambda(1 - \lambda)\|x - y\|^2 \\ \Rightarrow \frac{f(y + \lambda(x - y)) - f(y)}{\lambda} &\leq f(x) - f(y) - \mu(1 - \lambda)\|x - y\|^2 \\ \Rightarrow \nabla f(y)^\top(x - y) &\leq f(x) - f(y) - \mu\|x - y\|^2 \quad (\text{Grenzübergang } \lambda \searrow 0), \end{aligned}$$

d.h., es gilt (2.8).

Im Fall (a) geht man genauso vor mit $\mu = 0 \Rightarrow$ (2.6).

Im Fall (b) gilt ebenfalls (2.6), da f als strikt konvexe Funktion insbesondere konvex ist. Setze $z := (x + y)/2$, also gilt $x - y = 2(z - y)$. Es ergibt sich

$$\nabla f(y)^\top(x - y) = 2 \nabla f(y)^\top(z - y) \stackrel{(2.6)}{\leq} 2(f(z) - f(y)).$$

Für $x \neq y$ ist außerdem $f(z) < \frac{1}{2}f(x) + \frac{1}{2}f(y)$. Zusammen:

$$\nabla f(y)^\top(x - y) < f(x) - f(y),$$

also (2.7). □

Satz 2.25 (Charakterisierung konvexer Funktionen über 2. Ableitungen)

Sei $C \subset \mathbb{R}^n$ offen und konvex sowie $f : C \rightarrow \mathbb{R}$ zweimal stetig differenzierbar. Dann gelten:

- (a) f ist konvex auf $C \Leftrightarrow \nabla^2 f(x)$ ist positiv semidefinit für alle $x \in C$.
- (b) Ist $\nabla^2 f(x)$ positiv definit für alle $x \in C$, so ist f strikt konvex auf C .
- (c) f ist gleichmäßig konvex auf $C \Leftrightarrow$ Es existiert $\mu > 0$, sodass

$$d^\top \nabla^2 f(x) d \geq \mu \|d\|^2 \quad \text{für alle } d \in \mathbb{R}^n. \quad (2.9)$$

Beweis: Satz 3.8 in Numerische Verfahren zur Lösung unrestringierter Optimierungsaufgaben. Christian Kanzow und Carl Geiger, Springer, Berlin 1999 □

Eine Verallgemeinerung von (2.9) auf Operatoren zwischen Banachräumen führt auf das Konzept der Koerzivität oder Elliptizität (siehe späteres Kapitel über infinite Optimierung im zweiten Teil der Vorlesung).

Die Konvexität einer Funktion impliziert schon gewisse Stetigkeits- und Differenzierbarkeitseigenschaften (siehe späteres Kapitel zur nichtglatten Optimierung).

Wir betrachten die konvexe Optimierungsaufgabe

$$\text{Minimiere } f(x) \quad \text{über } x \in C \quad (2.10)$$

mit konvexer Grundmenge $C \subset \mathbb{R}^n$ und konvexer Zielfunktion $f : C \rightarrow \mathbb{R}$. [Beispiele für (2.10) sind Lineare Optimierungsprobleme, s. Paragraph 5].

Die fundamentale Bedeutung der Konvexität in der Optimierung erläutert

Satz 2.26 (Lokales und globales Optimum)

- (a) Jedes lokale Optimum von (2.10) ist bereits ein globales Optimum.
 (b) Die Lösungsmenge von (2.10) ist konvex (evtl. leer).
 (c) Ist f strikt konvex auf C , so besitzt (2.10) höchstens eine Lösung.

Beweis: (a): Sei x^* lokales Optimum, d.h., es existiert eine Umgebung $U(x^*)$ mit $f(x^*) \leq f(x)$ für alle $x \in C \cap U(x^*)$, vgl. Definition 2.11. Angenommen, es gäbe ein $\hat{x} \in C$ mit $f(\hat{x}) < f(x^*)$. Für $\lambda \in (0, 1]$ gilt dann

$$f(\lambda \hat{x} + (1 - \lambda)x^*) \leq \lambda f(\hat{x}) + (1 - \lambda)f(x^*) < f(x^*).$$

Für λ hinreichend klein liegt aber $\lambda \hat{x} + (1 - \lambda)x^* \in C \cap U(x^*)$, im Widerspruch zur lokalen Optimalität von x^* . Also kann \hat{x} nicht existieren, d.h., es ist

$$f(x^*) \leq f(x) \quad \text{für alle } x \in C.$$

(b): Seien x^* und x^{**} Lösungen von (2.10), also $f(x^*) = f(x^{**}) = f^* := \inf\{f(x) : x \in C\}$ (Optimalwert). Für $\lambda \in [0, 1]$ gilt

$$f(\lambda x^* + (1 - \lambda)x^{**}) \leq \lambda f(x^*) + (1 - \lambda)f(x^{**}) = f^*.$$

Da $\lambda x^* + (1 - \lambda)x^{**} \in C$ liegt, muss Gleichheit gelten, d.h., auch $\lambda x^* + (1 - \lambda)x^{**}$ ist Optimum von (2.10).

(c): Seien x^* und x^{**} zwei verschiedene Optima von (2.10). Wähle $\lambda = 1/2$:

$$\Rightarrow f\left(\frac{x^* + x^{**}}{2}\right) < \frac{1}{2}f(x^*) + \frac{1}{2}f(x^{**}) = f^*,$$

im Widerspruch zur Optimalität von x^* und x^{**} , da $\frac{x^* + x^{**}}{2} \in C$ liegt.

Definition 2.27

Zwei Mengen $K_1, K_2 \subset \mathbb{R}^n$ heißen **trennbar**, wenn es $\lambda \in \mathbb{R}^n \setminus \{0\}$ gibt, mit

$$\lambda^\top x \leq \alpha \leq \lambda^\top y \quad \forall x \in K_1, y \in K_2.$$

Die Hyperebene $H := \{x \in \mathbb{R}^n | \lambda^\top x = \alpha\}$ heißt **trennende Hyperebene**.

Satz 2.28 (Trennungssatz im \mathbb{R}^n)

Sei $K \subset \mathbb{R}^n$ konvex, $K \neq \emptyset$ und $y \notin \bar{K}$. Dann sind $\{y\}$ und K **trennbar**, d.h. $\exists \lambda \in \mathbb{R}^n \setminus \{0\}$ mit $\lambda^\top y \leq \lambda^\top x \quad \forall x \in K$. Ist $y \notin \bar{K}$, dann sind $\{y\}$ und K **echt trennbar**, d.h. $\exists \lambda \in \mathbb{R}^n \setminus \{0\}$ mit $\lambda^\top y < \lambda^\top x \quad \forall x \in K$.

Beweis:

1. Fall: $y \notin \bar{K}$.

Wähle die Norm $\|\cdot\| = \|\cdot\|_2$. Sei $d := \inf_{x \in \bar{K}} \|y - x\| > 0$.

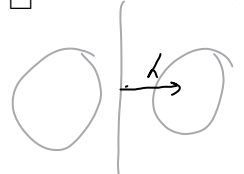
Die Funktion $f(x) := \|y - x\|$ nimmt auf dem Kompaktum $\bar{K} \cap \{x \in \mathbb{R}^n | \|x - y\| \leq 2d\}$ nach dem Satz von Weierstrass (Satz 2.3) ihr Minimum in einem Punkt $x_0 \in \bar{K}$ an. Definiere $\lambda := (x_0 - y)$.

Behauptung: λ erfüllt den Trennungssatz.

konvexe
Fkt

darf nicht
wieder
runter
gehen

⊥ orthogonal
zur Trennungslinie



$$\begin{aligned} \text{Rand} \quad \overline{(0,1)} &= [0,1] \\ \text{innere} \quad \overset{\circ}{(0,1)} &= (0,1) \end{aligned}$$

Sei $x \in \overline{K}$ beliebig. Da K konvex ist, gilt: $x_0 + s(x - x_0) \in \overline{K}$, $0 \leq s \leq 1$. Sei $s \neq 0$. Es gilt:

$$\begin{aligned} \|x_0 + s(x - x_0) - y\|^2 &\geq \|x_0 - y\|^2 \Leftrightarrow \langle x_0 + s(x - x_0) - y, x_0 + s(x - x_0) - y \rangle \\ &= \langle x_0 - y, x_0 - y \rangle + 2s\langle x_0 - y, x - x_0 \rangle + s^2\langle x - x_0, x - x_0 \rangle \\ &= \|x_0 - y\|^2 + 2s\langle x_0 - y, x - x_0 \rangle + s^2\|x - x_0\|^2 \geq \|x_0 - y\|^2 \\ &\Leftrightarrow 2s\langle x_0 - y, x - x_0 \rangle + s^2\|x - x_0\|^2 \geq 0 \end{aligned}$$

Division durch $2s > 0$ und Grenzübergang $s \rightarrow 0$ ergibt ...

$$\langle x_0 - y, x - x_0 \rangle = \langle \lambda, x - x_0 \rangle \geq 0 \Leftrightarrow \lambda^\top x \geq \lambda^\top x_0$$

Wegen $\|\lambda\| = d$ folgt:

$\lambda^\top x \geq \lambda^\top x_0 = \lambda^\top(x_0 - y + y) = \lambda^\top(x_0 - y) + \lambda^\top y = \|\lambda\|^2 + \lambda^\top y$ und da $\|\lambda\|^2 > 0$ ist $\lambda^\top x > \lambda^\top y$. Also ist $H = \{z \in \mathbb{R}^n \mid \lambda^\top z = \lambda^\top x_0 = \alpha\}$ trennende Hyperebene von $\{y\}$ und K .

2.Fall: $y \in \overline{K}$, also $y \in \partial K = K \setminus \overset{\circ}{K}$.

Zu $y \in \partial K$ Folge $\{y_k\}$, $y_k \notin \overset{\circ}{K}$ mit $\lim_{k \rightarrow \infty} y_k = y$. Für jedes dieser y_k gilt der 1.Fall
 $\Rightarrow \exists \lambda_k \in \mathbb{R}^n \setminus \{0\}$ mit $\lambda_k y_k < \lambda_k x \quad \forall x \in K$.

Normiere λ_k : $\frac{\lambda_k}{\|\lambda_k\|} y_k < \frac{\lambda_k}{\|\lambda_k\|} x \quad \forall x \in K$

$\frac{\lambda_k}{\|\lambda_k\|}$ ist Element der „Einheitssphäre“ E im \mathbb{R}^n und da $E \subset \mathbb{R}^n$ kompakt, $\exists \lambda \in \mathbb{R}^n$ (o.B.d.A. ggf. nach Übergang zu einer Teilfolge) mit $\lim_{k \rightarrow \infty} \lambda_k = \lambda$ und es gilt $\lambda y \leq \lambda x$ für eben dieses λ .

□

Bemerkung 2.29 (Verallgemeinerung auf ∞ -dimensionale Räume)

Da in ∞ -dimensionalen Räumen beschränkte und abgeschlossene Mengen nicht notwendig kompakt sind, geht die obige Beweiskette in diesem Fall schief. Es werden zusätzlich funktionalanalytische Konzepte benötigt, mit deren Hilfe dann allerdings dieselben abstrakten Prinzipien wie im obigen Beweis zum Einsatz kommen. Das führt auf den Satz von Hahn-Banach in Trennungssatz-Form, der das ∞ -dimensionale Analogon zum Trennungssatz (siehe Lehrbücher der Funktionalanalysis, z.B. Dirk Werner: Funktionalanalysis, Springer 2007) darstellt.

Korollar 2.30

Sei $K \subset \mathbb{R}^n$ ein nicht-leerer, konvexer und abgeschlossener Kegel. Dann gibt es zu $y \notin K$ ein $\lambda \in \mathbb{R}^n \setminus \{0\}$ mit $\lambda y < 0 \leq \lambda x \quad \forall x \in K$.

Beweis: Übungsaufgabe

□

Lemma 2.31 (Lemma von Farkas)

Sei B eine beliebige $(k \times n)$ -Matrix und $d \in \mathbb{R}^k$. Dann gilt genau eine der beiden Aussagen

- i) $Bx = d, x \geq 0$ hat eine Lösung $x \in \mathbb{R}^n$
- ii) $\lambda B \geq 0, \lambda d < 0$ hat eine Lösung $\lambda \in \mathbb{R}^k$

Beweis: Der Kegel $K := \{Bx \mid x \geq 0, x \in \mathbb{R}^n\} \subset \mathbb{R}^k$ ist nicht-leer, konvex und abgeschlossen.

Es gilt genau eine der beiden Aussagen

- a) $d \in K$, d.h. $\exists x \geq 0, x \in \mathbb{R}^n$ mit $Bx = d$
- b) $d \notin K$

Zeige, dass aus b) Aussage ii) des Lemmas folgt.

Für $d \notin K$ folgt mit Korollar 2.30: $\exists \lambda \in \mathbb{R}^k \setminus \{0\}$ mit $\lambda d < 0 \leq \lambda z \forall z \in K$

$\Rightarrow 0 \leq \lambda(Bx) = (\lambda B)x, x \geq 0, x \in \mathbb{R}^n$

$\Rightarrow (\lambda B) \geq 0$, da die vorherige Aussage für alle $x \in \mathbb{R}^n, x \geq 0$ gelten muss.

□

Aufgrund seiner exklusiven “ODER”-Aussage in i) und ii) nennt man das *Farkas-Lemma* auch **Alternativsatz von Farkas**. Es wird in vielen Lehrbüchern zum Beweis der notwendigen Optimalitätsbedingungen erster Ordnung für beschränkte Optimierungsprobleme verwendet. In dieser Vorlesung verwenden wir dafür den Trennungssatz direkt.

Definition 2.32 (Extremalpunkt/Ecke)

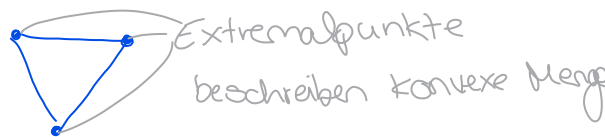
Sei $K \subset \mathbb{R}^n$ konvex. Ein Punkt $x \in K$ heißt Extremalpunkt (oder Ecke) von K , wenn es **nicht** zwei Punkte $y, z \in K, x \neq y, x \neq z$ gibt mit $x = \alpha y + (1 - \alpha)z$ für ein $\alpha \in (0, 1)$ oder alternativ gilt:

$$x = \alpha y + (1 - \alpha)z, y \in K, z \in K, \alpha \in (0, 1) \Rightarrow x = y = z.$$

Satz 2.33

Eine kompakte und konvexe Menge $K \subset \mathbb{R}^n$ ist der Abschluss der konvexen Hülle ihrer Extremalpunkte.

Beweis: Übungsaufgabe.



□

Bemerkung 2.34

Die Verallgemeinerung dieses Resultats auf den ∞ -dimensionalen Fall liefert den Satz von Krein-Milman. Sein Beweis verwendet zwei tiefliegende Konzepte, den Satz von Hahn-Banach und das Zorn'sche Lemma.

§ 3 Nichtlineare Optimierung

§ 3.1 Notwendige Optimalitätsbedingungen

Zur Erinnerung: Notwendige Bedingungen im **unbeschränkten Fall**

$$f : \mathbb{R}^n \rightarrow \mathbb{R}, \bar{x} \text{ lokales Minimum von } f \Rightarrow \nabla f(\bar{x}) = 0$$

Hinreichende Bedingung $\nabla f(\bar{x}) = 0$ und $\nabla^2 f(\bar{x})$ positiv definit.

Übertrage notwendige Bedingungen auf den beschränkten Fall.

Im Falle von Beschränkungen sind die o.a. Bedingungen nicht mehr sinnvoll und vor allem nicht mehr notwendig, da nicht alle "Richtungen" im Raum zulässig sind, wenn man sich am Rand der zulässigen Menge befindet.

Ziel: Herleitung notwendiger Optimalitätsbedingungen über eine Charakterisierung „der zulässigen Richtungen“!

Zentrale Konzepte:

1. Satz über implizite Funktionen
2. Trennungssatz für konvexe Mengen

Mit Hilfe von 1. + 2. lassen sich konzeptionell notwendige Optimalitätsbedingungen für den beschränkten Fall ableiten.

Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$ und $S \subset \mathbb{R}^n$, $\min f(x)$ s.t. $x \in S$.

Minimalannahme für die Theorie: $S = \{x \in \mathbb{R}^n \mid g(x) \in K\}$ mit konvexer Menge K und sog. Restriktionsabbildung g .

Vorteil: „Koordinatenfreie“ Formulierung (auch geeignet für ∞ -dim. Banachräume)

In beliebigen Räumen Y induziert ein konvexer Kegel K (Ordnungskegel) eine Halbordnung durch $y \leq 0 \Leftrightarrow y \in K$!

Für den Spezialfall $K \subset \mathbb{R}^k, K \times \{0\}_{m-k} \subset \mathbb{R}^m$, zerlege die Abbildung g in $g = (g_1, \dots, g_k)$ und $h = (g_{k+1}, \dots, g_m)$.

Optimierungsproblem (NLP):

$$\min \{f(x) \mid g(x) \in K, h(x) = 0\}, \text{ wähle } K = \mathbb{R}_-^k = \{y \in \mathbb{R}^k \mid y_i \leq 0, i = 1, \dots, k\}$$

\rightarrow Standardproblem der Nichtlinearen Optimierung im \mathbb{R}^n

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & g_i(x) \leq 0, \quad i = 1, \dots, k \\ & g_i(x) = 0, \quad i = k+1, \dots, m \end{aligned}$$

Annahme im folgenden: Seien f, g, h hinreichend oft stetig differenzierbar.

Sei \bar{x} zulässiger Punkt von NLP, definiere Indexmengen

$$I(\bar{x}) := \{i \in \{1, \dots, k\} \mid g_i(\bar{x}) = 0\} \quad (\text{aktive Ungleichungen})$$

$$J(\bar{x}) := I(\bar{x}) \cup \{k+1, \dots, m\} \quad (\text{aktive Ungleichungen} + \text{Gleichungen})$$

Definition 3.1

Der **Tangentialkegel** einer Menge S bezüglich $\bar{x} \in S$ ist definiert als:

$$T(S, \bar{x}) := \left\{ v \in \mathbb{R}^n : \exists \text{ Punkte } x_i \in S \text{ und } t_i \in \mathbb{R}_+ \text{ mit } \lim_{i \rightarrow \infty} t_i = 0, v = \lim_{i \rightarrow \infty} \frac{x_i - \bar{x}}{t_i} \right\}$$

D.h. es gilt für Tangentialkegelelemente die Darstellung

$$v \in T(S, \bar{x}) : x_i = \bar{x} + t_i v + r_i, r_i = o(t_i), \text{ d.h. } \lim_{i \rightarrow \infty} \frac{r_i}{t_i} = 0$$

Lemma 3.2

Für jede Menge $S \subset \mathbb{R}^n$ ist $T(S, \bar{x})$ ein abgeschlossener Kegel (mit Spitze in 0).

Beweis: Übungsaufgabe □

Lemma 3.3

Sei $K \subset \mathbb{R}^n$ konvex und sei $\bar{x} \in K$. Der Tangentialkegel $T(K, \bar{x})$ ist der Abschluss der konischen Hülle von K in \bar{x} , d.h. $\overline{K(\bar{x})} = T(K, \bar{x})$.

Beweis: Übungsaufgabe □

Satz 3.4 (Variationsungleichung)

Sei $S \subset \mathbb{R}^n$, $\bar{x} \in S$ eine lokale Minimalstelle des Problems $\min \{f(x) \mid x \in S\}$. Sei f differenzierbar in \bar{x} , dann gilt:

$$\nabla f(\bar{x})^\top v \geq 0 \quad \forall v \in T(S, \bar{x})$$

Beweis: Übungsaufgabe □

Satz 3.4 liefert eine notwendige Optimalitätsbedingung ! Sei im folgenden $S \subset \mathbb{R}^n$ gegeben durch $S = \{x \in \mathbb{R}^n \mid g(x) \in K\}$, $K \subset \mathbb{R}^m$ konvex.

Der Tangentialkegel (nach Def. 3.1) ist i.a. in der praktischen Anwendung schwer zu handhaben und damit lässt sich Satz 3.4 nicht ohne weiteres anwenden.

Ziel: Charakterisiere $T(S, \bar{x})$ mit Hilfe der Restriktionsabbildung g !

Motivation: Linearisiere die Inklusion $g(x) \in K$

$$\rightarrow g(\bar{x}) + g'(\bar{x})v \in K, \quad v \in \mathbb{R}^n$$

$$\Rightarrow g'(\bar{x})v \in K - g(\bar{x}) \subset \bigcup_{\alpha > 0} \alpha(K - g(\bar{x})) = K(g(\bar{x}))$$

Definition 3.5

Der **Linearisierende Kegel** der Menge S bezüglich $\bar{x} \in S$ ist definiert als:

$$L(S, \bar{x}) := \{v \in \mathbb{R}^n \mid g'(\bar{x})v \in K(g(\bar{x}))\}$$

Beispiel 3.6 (Standardbeispiel)

Kegel $K = \mathbb{R}_-^k \times \{0\}_{m-k}$

$$S = \{x \in \mathbb{R}^n \mid g_i(x) \leq 0, \ i = 1, \dots, k \\ g_i(x) = 0, \ i = k+1, \dots, m\}$$

$$K(g(\bar{x})) = \{y \in \mathbb{R}^m \mid y_i \leq 0, \text{ falls } i \in I(\bar{x}) \\ y_i = 0, \text{ falls } i = k+1, \dots, m\}$$

Linearisierender Kegel

$$L(S, \bar{x}) = \{v \in \mathbb{R}^n \mid g'_i(\bar{x})v \leq 0, \ i \in I(\bar{x}) \\ g'_i(\bar{x})v = 0, \ i = k+1, \dots, m\}$$

Lemma 3.7

Ist $K(g(\bar{x}))$ abgeschlossen, so gilt $T(S, \bar{x}) \subset L(S, \bar{x})$.

Beweis: Übungsaufgabe □

Der Nachweis von $L(S, \bar{x}) \subset T(S, \bar{x})$ ist deutlich schwieriger und erfordert eine sog. Regularisierungsbedingung (**CQ, constraint qualification**).

Definition 3.8 (Abadie-CQ)

Ein Punkt $\bar{x} \in S$ genügt der **Regularitätsbedingung von Abadie** (*Abadie constraint qualification* oder *Abadie-CQ*), wenn $T(S, \bar{x}) = L(S, \bar{x})$ gilt. ◇

Die *Abadie-CQ* ist in der Praxis schwierig nachprüfbar, es gibt aber einfacher überprüfbare (stärkere) **CQ**, welche die *Abadie-CQ* implizieren. Betrachten wir zunächst den gleichungsbeschränkten Fall, sei also zunächst $S = \{x \in \mathbb{R}^n \mid g(x) = 0\}$.

Definition 3.9

Im rein **gleichungsbeschränkten** Fall heißt ein Punkt $\bar{x} \in S$ **regulär**, wenn gilt: Die Gradienten $g'_i(\bar{x})$ sind **linear unabhängig** für $i = 1, \dots, m$ bzw. $\text{Im}(g'(\bar{x})) = \mathbb{R}^m$ (also $g'(\bar{x})$ ist surjektiv). *Alternative Bedingung (für den Fall $m < n$, siehe auch Beweis von Satz 5.20):* Die Matrix $g'(\bar{x})g'(\bar{x})^\top$ ist regulär (Übungsaufgabe!).

Bemerkung 3.10

Es gilt genau eine der beiden Aussagen:

- (i) $\bar{x} \in S$ ist regulär
- (ii) $\exists \lambda \in \mathbb{R}^m, \lambda \neq 0$ mit $\lambda g'(\bar{x}) = 0$

Satz 3.11

Sei $S = \{x \in \mathbb{R}^n \mid g(x) = 0\}$ und sei $\bar{x} \in S$ regulär. Dann gilt:

- (i) Zu $v \in \mathbb{R}^n$ mit $g'(\bar{x})v = 0$ (d.h. $v \in L(S, \bar{x})$) gibt es $\epsilon > 0$ und eine Kurve $x : [-\epsilon, \epsilon] \rightarrow S$ mit $x(0) = \bar{x}$ und $\lim_{t \rightarrow 0} \frac{x(t) - \bar{x}}{t} = \dot{x}(0) = v$.
- (ii) $T(S, \bar{x}) = L(S, \bar{x})$.

Beweis: Definiere $F : \mathbb{R}^{m+1} \rightarrow \mathbb{R}^m$, $F(y, t) := g(\bar{x} + tv + g'(\bar{x})^\top y)$

Ziel: Bestimme y als Funktion $y(t)$ mit $F(y(t), t) = 0$ für $t \in [-\epsilon, \epsilon]$. Nach Voraussetzung gilt: $F(0, 0) = 0$, $\frac{\partial F}{\partial y}(0, 0) = g'(\bar{x})g'(\bar{x})^\top$ ist regulär (da \bar{x} regulär).

Nach dem Satz über implizite Funktionen gibt es $\epsilon > 0$ und eine in $t = 0$ stetig differenzierbare Kurve $y(t)$, $t \in [-\epsilon, \epsilon]$ mit $y(0) = 0$, $F(y(t), t) = 0 \forall t \in [-\epsilon, \epsilon]$.

$$\Rightarrow 0 = \frac{d}{dt} F(y(t), t) |_{t=0} = g'(\bar{x})g'(\bar{x})^\top \dot{y}(0) + g'(\bar{x})v = g'(\bar{x})g'(\bar{x})^\top \dot{y}(0)$$

$$g'(\bar{x})g'(\bar{x})^\top \text{ regulär} \Rightarrow \dot{y}(0) = 0$$

$$\dot{y}(0) = \lim_{t \rightarrow 0} \frac{y(t) - y(0)}{t} = \lim_{t \rightarrow 0} \frac{y(t)}{t}$$

Definiere $x(t) := \bar{x} + tv + g'(\bar{x})^\top y(t)$, $t \in [-\epsilon, \epsilon]$. Dann gilt: $g(x(t)) = 0$, $t \in [-\epsilon, \epsilon]$

$$\Rightarrow x(t) \in S, \quad x(0) = \bar{x}, \quad \dot{x}(0) = v \quad (\text{wegen } \dot{y}(0) = 0) \quad \square$$

Wir betrachten nun den allgemeinen Fall. Sei dazu $S = \{x \in \mathbb{R}^n \mid g(x) \in K\}$, $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $K \subset \mathbb{R}^m$ konvex.

Definition 3.12 (Regularität nach Robinson, 1975 (auch Maurer, Zowe, Kurcyusz, 1979))

Im allgemeinen Fall ($S = \{x \in \mathbb{R}^n \mid g(x) \in K \subset \mathbb{R}^m\}$) heißt ein Punkt $\bar{x} \in S$ **(Robinson-)regulär**, falls $\text{Im}(g'(\bar{x})) - K(g(\bar{x})) = \mathbb{R}^m$, d.h. $\text{Im}(g'(\bar{x}))$ und $K(g(\bar{x}))$ sind "transversal" zueinander ("Transversalitätsbedingung").

Lemma 3.13

Die folgenden Bedingungen sind äquivalent:

- i) $\text{Im}(g'(\bar{x})) - K(g(\bar{x})) = \mathbb{R}^m$
- ii) $0 \in \text{int}(\text{Im}(g'(\bar{x})) - K(g(\bar{x})))$
- iii) $0 \in \text{int}(\text{Im}(g'(\bar{x})) + g(\bar{x}) - K)$

Beweis: $i) \Rightarrow iii) \Rightarrow ii) \Rightarrow i)$, hier nur $i) \Rightarrow iii)$, der Rest: Übungsaufgabe

Die Menge $(\text{Im}(g'(\bar{x})) + g(\bar{x}) - K)$ ist konvex und enthält die 0 wegen $g(\bar{x}) \in K$. Angenommen $0 \notin \text{int}(\text{Im}(g'(\bar{x})) + g(\bar{x}) - K)$. Nach dem Trennungssatz gibt es $\lambda \in \mathbb{R}^m$, $\lambda \neq 0$ mit $\lambda^\top (g'(\bar{x})v + g(\bar{x}) - y) \geq 0 \forall v \in \mathbb{R}^n, y \in K$

Da $K(g(\bar{x})) = \bigcup_{\alpha > 0} \alpha(K - g(\bar{x}))$ folgt (Multiplikation mit $\alpha > 0$):

$$\lambda^\top (g'(\bar{x})v - y) \geq 0 \quad \forall y \in K(g(\bar{x})), v \in \mathbb{R}^n \quad (3.1)$$

Die Voraussetzung $\text{Im}(g'(\bar{x})) - K(g(\bar{x})) = \mathbb{R}^m$ impliziert $\lambda = 0$ im Widerspruch zum Trennungssatz. \square

Bemerkung 3.14

Ein Punkt $\bar{x} \in S$ ist nicht regulär $\Leftrightarrow \exists \lambda \in \mathbb{R}^m, \lambda \neq 0$ mit $\lambda g'(\bar{x}) = 0, \lambda(-y) \geq 0 \forall y \in K(g(\bar{x}))$. Das folgt aus (3.1).

Spezialisierung für den Fall $S = \{x \in \mathbb{R}^n : g(x) \in K, h(x) = 0\}$ mit $K \subset \mathbb{R}^k, K^\circ \neq \emptyset, g = (g_1, \dots, g_k), h = (g_{k+1}, \dots, g_m)$:

Bestätige durch Nachrechnen, dass Bedingung ii) in Lemma 3.13 äquivalent ist zu

$$\begin{aligned} \text{Im}(h'(\bar{x})) = \mathbb{R}^{m-k}, \text{ d.h. } g'_i(\bar{x}), i = k+1, \dots, m \text{ sind linear unabhängig} \\ \text{und } \exists v \in \mathbb{R}^n \text{ mit } h'(\bar{x})v = 0, g'(\bar{x})v \in \text{int}(K(g(\bar{x}))) \end{aligned} \quad (3.2)$$

Bedingung iii) in Lemma 3.13 ist äquivalent zu

$$\begin{aligned} \text{Im}(h'(\bar{x})) = \mathbb{R}^{m-k}, \text{ d.h. } g'_i(\bar{x}), i = k+1, \dots, m \text{ sind linear unabhängig} \\ \text{und } \exists v \in \mathbb{R}^n \text{ mit } h'(\bar{x})v = 0 \\ g(\bar{x}) + g'(\bar{x})v \in \overset{\circ}{K} \end{aligned} \quad (3.3)$$

Beispiel 3.15 (Standardfall $K = \mathbb{R}_-^k$)

$$\begin{aligned} S = \{x \in \mathbb{R}^n \mid g_i(x) \leq 0, i = 1, \dots, k \quad \text{ungleichheits Bed.} \\ g_i(x) = 0, i = k+1, \dots, m\} \quad \text{gleichheits Bed.} \end{aligned}$$

Die Transversalitätsbedingung aus Def. 3.12 heißt in diesem Fall *Mangasarian-Fromowitz-CQ (MFCQ)* und lautet:

Die Gradienten $g'_{k+1}(\bar{x}), \dots, g'_m(\bar{x})$ sind linear unabhängig und es gibt $v \in \mathbb{R}^n$ mit $g'_i(\bar{x})v < 0 \forall i \in I(\bar{x})$ und $g'_i(\bar{x})v = 0$ für $i = k+1, \dots, m$

Lemma 3.16

Sei $K \subset \mathbb{R}^n$ konvex, $x \in \bar{K}, y \in \overset{\circ}{K}$. Dann gilt $z = \lambda x + (1 - \lambda)y \in \overset{\circ}{K}, \lambda \in [0, 1)$.

Beweis: Siehe Lemma 2.20. □

Satz 3.17

Sei $\bar{x} \in \mathbb{R}^n$ ein regulärer Punkt der zulässigen Menge $S = \{x \in \mathbb{R}^n \mid g(x) \in K, h(x) = 0\}, K \subset \mathbb{R}^k, K$ konvex, $\overset{\circ}{K} \neq \emptyset$. Dann gilt

- (i) Zu $v_0 \in \mathbb{R}^n$ mit $h'(\bar{x})v_0 = 0$ und $g(\bar{x}) + g'(\bar{x})v_0 \in \overset{\circ}{K} \exists \epsilon > 0$ und eine Kurve $x : [0, \epsilon] \rightarrow S$ mit $x(0) = \bar{x}$ und $v_0 = \lim_{t \searrow 0} \frac{x(t) - \bar{x}}{t} = \dot{x}(0)$
- (ii) Es gilt $L(S, \bar{x}) \subset T(S, \bar{x})$

Beweis:

Zu (i): Im Falle $k = m$ (keine Gleichungsbeschränkungen) setze man $x(t) = \bar{x} + tv_0$ für $t \geq 0$.

Falls $k < m$ (es treten Gleichungsbeschränkungen auf), so existiert nach Satz 3.11 (wegen $h'(\bar{x})v_0 = 0$) eine Zahl $\delta > 0$ und eine Kurve $x : [-\delta, \delta] \rightarrow \mathbb{R}^n$ mit $h(x(t)) = 0 \forall t \in [-\delta, \delta], x(0) = \bar{x}, v_0 = \lim_{t \searrow 0} \frac{x(t) - \bar{x}}{t} = \dot{x}(0)$.

Wegen $g(\bar{x}) + g'(\bar{x})v_0 \in \overset{\circ}{K}$ und $\lim_{t \rightarrow 0} \frac{g(x(t)) - g(\bar{x})}{t} = g'(\bar{x})\dot{x}(0) = g'(\bar{x})v_0$

folgt $g(\bar{x}) + \frac{g(x(t)) - g(\bar{x})}{t} \in K$ für $t \in [-\epsilon, \epsilon]$ für geeignetes $\epsilon \leq \min\{1, \delta\}$.

Wegen $g(\bar{x}) \in K$ und der Konvexität von K folgt:

$g(x(t)) = (1 - t)g(\bar{x}) + t \left(g(\bar{x}) + \frac{g(x(t)) - g(\bar{x})}{t} \right) \in K$ für $t \in [0, \epsilon]$ mit $\epsilon > 0$ geeignet.

Also gilt $g(x(t)) \in K$, d.h. $x(t) \in S$ für $0 \leq t \leq \epsilon$ (daher $\epsilon \leq 1$ erforderlich ! \rightarrow

Konvexkombination).

Zu (ii): Es gilt $L(S, \bar{x}) = \{v \in \mathbb{R}^n \mid g'(\bar{x})v \in K(g(\bar{x})), h'(\bar{x})v = 0\}$. Sei $v \in L(S, \bar{x})$ gegeben. Nach Definition von $K(g(\bar{x})) \exists r > 0$ mit $g'(\bar{x})v \in r(K - g(\bar{x})), h'(\bar{x})v = 0$.
 $\Rightarrow g(\bar{x}) + g'(\bar{x})\frac{v}{r} \in K, h'(\bar{x})v = 0$

Wegen der Regularität von $\bar{x} \in S$ gibt es $v_0 \in \mathbb{R}^n$ mit $g(\bar{x}) + g'(\bar{x})v_0 \in \overset{\circ}{K}, h'(\bar{x})v_0 = 0$
 Bilde Konvexkombination $v_\alpha := (1 - \alpha)v_0 + \alpha\frac{v}{r}, 0 \leq \alpha \leq 1$ Nach Lemma 3.16 gilt:
 $(1 - \alpha)(g(\bar{x}) + g'(\bar{x})\frac{v}{r}) + \alpha(g(\bar{x}) + g'(\bar{x})v_0) = g(\bar{x}) + g'(\bar{x})v_\alpha \in \overset{\circ}{K}$ für $0 \leq \alpha < 1$ (!)
 Nach (i) folgt dann $v_\alpha \in T(S, \bar{x})$ $0 \leq \alpha < 1$ (da v_α die Voraussetzungen von (i) erfüllt). Wegen der Abgeschlossenheit von $T(S, \bar{x})$ folgt: $v = \lim_{\alpha \rightarrow 1} rv_\alpha \in T(S, \bar{x})$
 $(rv_\alpha \in T(S, \bar{x})$ aufgrund der Kegeleigenschaft). Es folgt also $L(S, \bar{x}) \subset T(S, \bar{x})$. \square

Beispiel 3.18 (Regularität)

Sei

$$S := \{x = (x_1, x_2) \in \mathbb{R}^2 : g_1(x) = x_2 - x_1^3 \leq 0, g_2(x) = x_2 \leq 0\}$$

und $\bar{x} = (0, 0)$. Dann ist $I(\bar{x}) = \{1, 2\}$ und die Gradienten $g'_1(\bar{x}) = g'_2(\bar{x}) = (0, 1)$ sind linear abhängig. Es gilt für jeden Vektor aus $v \in \mathbb{R}^2$ mit $v_2 < 0 : g'_i(\bar{x})v = v_2 < 0$ für $i = 1, 2$ und weiter ist

$$T(S, \bar{x}) = L(S, \bar{x}) = \{v \in \mathbb{R}^2 : v_2 \leq 0\}.$$

Demnach ist \bar{x} regulär im Sinne von Abadie. Ändere nun S ab zu

$$S := \{x = (x_1, x_2) \in \mathbb{R}^2 : g_1(x) = x_2 - x_1^3 \leq 0, g_2(x) = -x_2 \leq 0\}$$

Dann ist \bar{x} **nicht** regulär, denn die Bedingungen

$$g'_1(\bar{x})v = v_2 < 0 \quad \text{und} \quad g'_2(\bar{x})v = -v_2 < 0$$

sind für kein $v \in \mathbb{R}^2$ erfüllbar. Hier gilt

$$\begin{aligned} T(S, \bar{x}) &= \{v \in \mathbb{R}^2 : v_1 \geq 0, v_2 = 0\} \\ L(S, \bar{x}) &= \{v \in \mathbb{R}^2 : g'_1(\bar{x})v = v_2 \leq 0, g'_2(\bar{x})v = -v_2 \leq 0\} \\ &= \{v \in \mathbb{R}^2 : v_2 = 0\}. \end{aligned}$$

Also folgt insgesamt $T(S, \bar{x}) \subsetneq L(S, \bar{x})$.

Mit Hilfe der Regularitätsresultate kann man nun an die Herleitung notwendiger Optimalitätsbedingungen gehen.

Optimierungsproblem:

$$\min \{f(x) \mid g(x) \in K\}, g : \mathbb{R}^n \rightarrow \mathbb{R}^m, K \subset \mathbb{R}^m \text{ konvex} \quad (3.4)$$

Voraussetzung: Sei $\bar{x} \in S = \{x \in \mathbb{R}^n \mid g(x) \in K\}$ eine lokale Minimalstelle. Seien f und g stetig differenzierbar in einer Umgebung von \bar{x} . Ideen:

- Linearisierung von f und g
- Variationsungleichung $f'(\bar{x})v \geq 0 \quad \forall v \in T(S, \bar{x})$
- $L(S, \bar{x}) = T(S, \bar{x})$ (Regularität von \bar{x}) $\rightarrow f'(\bar{x})v \geq 0 \quad \forall v \in L(S, \bar{x})$
- Trennungssatz für konvexe Mengen (!)

Motivation:

Es gibt eine Umgebung U von \bar{x} mit

$\{(f(x) - f(\bar{x}), g(x)) \in \mathbb{R} \times \mathbb{R}^m \mid x \in U\} \cap \{(r, y) \in \mathbb{R} \times \mathbb{R}^m \mid r < 0, y \in K\} = \emptyset$,
wegen $g(x) \in K \Rightarrow f(x) - f(\bar{x}) \geq 0$, da \bar{x} lokale Minimalstelle.

Daher ist also $(0, 0) \in \mathbb{R} \times \mathbb{R}^m$ ein Randpunkt der Differenzmenge

$$B := \{(f(x) - f(\bar{x}) + r, g(x) - y) \mid x \in U, r \geq 0, y \in K\}$$

Allerdings ist B i.a. nicht konvex.

Daher: Konvexifizierung von B durch Linearisierung (Taylor-Entwicklung 1. Ordnung um \bar{x}):

$$B \rightarrow \tilde{B} = \{(f'(\bar{x})v + r, g'(\bar{x})v + g(\bar{x}) - y) \mid v \in \mathbb{R}^n, r \geq 0, y \in K\}$$

Betrachte die konische Hülle von \tilde{B} bzgl. 0:

$$A := \{(f'(\bar{x})v + r, g'(\bar{x})v - y) \mid v \in \mathbb{R}^n, r \geq 0, y \in K(g(\bar{x}))\} \quad (3.5)$$

A ist konvexe Approximation (konvexer Kegel) der i.a. nicht konvexen Menge B . Die Aussage des Satzes von Karush-Kuhn-Tucker ist äquivalent zu der Aussage $(0, 0) \in \mathbb{R} \times \mathbb{R}^m$ ist ein Randpunkt von A .

Satz 3.19 (Satz von John, Karush, Kuhn, Tucker)

Sei $\bar{x} \in S$ lokale Minimalstelle von $\min_{x \in S} f(x)$.

(i) **(Fritz John-Bedingung)**

Es gibt einen „Zeilenvektor“ $(\lambda_0, \lambda) \in \mathbb{R} \times \mathbb{R}^m$, $(\lambda_0, \lambda) \neq (0, 0)$ mit

$$\lambda_0 f'(\bar{x}) + \lambda g'(\bar{x}) = 0, \lambda_0 \geq 0, \lambda(-y) \geq 0 \forall y \in K(g(\bar{x}))$$

(ii) **(KKT-Bedingungen, Karush-Kuhn-Tucker)**

Ist \bar{x} regulär, so gilt $\lambda_0 > 0$ in (i). Daher gibt es $\lambda \in \mathbb{R}^m$ mit

$$f'(\bar{x}) + \lambda g'(\bar{x}) = 0, \lambda(-y) \geq 0 \forall y \in K(g(\bar{x})),$$

$$\text{d.h. für } K := \mathbb{R}_+^k \times \{0_{m-k}\} : \lambda_i \geq 0 \forall i \in I(\bar{x})$$

Beweis: (i), (ii) gemeinsam.

1. Fall: \bar{x} sei nicht regulär. Nach Bemerkung 3.14 \exists dann $\lambda \in \mathbb{R}^m$, $\lambda \neq 0_m$ mit $\lambda g'(\bar{x}) = 0$, $\lambda(-y) \geq 0 \forall y \in K(g(\bar{x}))$. Dann ist Beh. (i) mit $\lambda_0 = 0$ erfüllt.

2. Fall: \bar{x} sei regulär. Die Variationsungleichung (Satz 3.4) lautet $f'(\bar{x})v \geq 0 \forall v \in T(S, \bar{x})$.

Mit Satz 3.17 (ii) folgt:

$$L(S, \bar{x}) = \{v \in \mathbb{R}^n \mid g'(\bar{x})v \in K(g(\bar{x}))\} \subset T(S, \bar{x})$$

Also:

$$f'(\bar{x})v \geq 0 \quad \forall v \in \mathbb{R}^n, g'(\bar{x})v \in K(g(\bar{x})). \quad (3.6)$$

Daher ist (siehe 3.5) $(0, 0_m)$ ein Randpunkt des konvexen Kegels

$$A = \{(f'(\bar{x})v + r, g'(\bar{x})v - y) \mid v \in \mathbb{R}^n, r \geq 0, y \in K(g(\bar{x}))\},$$

Nach dem Trennungssatz 2.28 \exists Zeilenvektor $(\lambda_0, \lambda) \in \mathbb{R} \times \mathbb{R}^m \neq (0, 0_m)$ mit

$$\lambda_0(f'(\bar{x})v + r) + \lambda(g'(\bar{x})v - y) \geq 0 \quad \forall v \in \mathbb{R}^n, r \geq 0, y \in K(g(\bar{x})). \quad (3.7)$$

Setze $r = 0$, $y = 0 \rightarrow (\lambda_0 f'(\bar{x}) + \lambda g'(\bar{x}))v \geq 0 \quad \forall v \in \mathbb{R}^n$

$$\implies \lambda_0 f'(\bar{x}) + \lambda g'(\bar{x}) = 0 \quad (!)$$

Setze $r = 0$, $v = 0 \rightarrow \lambda(-y) \geq 0 \quad \forall y \in K(g(\bar{x}))$. Setze $v = 0, y = 0, r = 1$, dann folgt $\lambda_0 \geq 0$. Damit ist (i) bewiesen. [Für f linear, d.h. $f(x) = cx$, liefert das gerade Dualität für LP !!! (Satz 5.13)]

Es gilt $\lambda_0 > 0$, da \bar{x} regulär, denn $\lambda_0 = 0 \Leftrightarrow \bar{x}$ nicht regulär (nach Bemerkung 3.14).

Übergang in (i) $\lambda \rightarrow \frac{\lambda}{\lambda_0}$, d.h. setze o.B.d.A. $\lambda_0 = 1$.

\implies Beh. (ii) □

Der Satz von Karush-Kuhn-Tucker heißt auch „Lagrange’sche Multiplikatoren“- Regel, die Komponenten λ_i , $i = 0, \dots, m$ Lagrange-Multiplikatoren.

Bemerkung 3.20

Der Lagrange-Multiplikator λ aus dem Satz von Karush-Kuhn-Tucker ist i.a. **nicht** eindeutig bestimmt !

Bemerkung 3.21

Ist die Menge K ein konvexer Kegel, so gilt $K(g(\bar{x})) = K + \mathbb{R} \cdot g(\bar{x})$ und die Aussage $[\lambda(-y) \geq 0 \quad \forall y \in K(g(\bar{x}))]$ ist äquivalent zu der Aussage

$$[\lambda(-y) \geq 0 \quad \forall y \in K \text{ und } \lambda g(\bar{x}) = 0] \quad (3.8)$$

Die Gleichung $\lambda g(\bar{x}) = 0$ heißt **Komplementaritätsbedingung**.

Beweis: (der Darstellung $K(g(\bar{x})) = K + \mathbb{R} \cdot g(\bar{x})$) Ein Element von $K(g(\bar{x}))$ hat die Darstellung $\alpha(y - g(\bar{x})), y \in K, \alpha > 0$, es genügt daher zu zeigen, dass mit $(y - g(\bar{x})), y \in K$ auch $(y + g(\bar{x})), y \in K$ Element von $K(g(\bar{x}))$ ist. Da K ein Kegel ist, gilt mit $g(\bar{x}) \in K$ auch $3g(\bar{x}) \in K$ und damit folgt $2g(\bar{x}) = 1 \cdot (3g(\bar{x}) - g(\bar{x})) \in K(g(\bar{x}))$. Bilde wegen der Konvexität von $K(g(\bar{x}))$ Konvexkombination $\beta(y - g(\bar{x})) + (1 - \beta)2g(\bar{x}) = \frac{y}{2} + \frac{g(\bar{x})}{2}$ für $y \in K, \beta = \frac{1}{2}$ Konvexkombination und multipliziere mit $\alpha = 2$. □

Mit der Definition der sog. **Lagrange-Funktion**

$$\mathcal{L}(x, \lambda_0, \lambda) = \lambda_0 f(x) + \lambda g(x) = \lambda_0 f(x) + \sum_{i=1}^m \lambda_i g_i(x)$$

lautet die Bedingung (i) aus dem KKT-Satz dann

$$\mathcal{L}_x(\bar{x}, \lambda_0, \lambda) = \lambda_0 f'(\bar{x}) + \lambda g'(\bar{x}) = 0, \quad \lambda_0 \geq 0, \quad \lambda(-y) \geq 0 \quad \forall y \in K(g(\bar{x})).$$

D.h. mit Hilfe der Lagrange-Funktion \mathcal{L} „reduziert“ man ein beschränktes Optimierungsproblem in geeigneter Art und Weise auf den unbeschränkten Fall bzw. auf sehr einfache Ungleichungsbeschränkungen.

Die Menge der Lagrange-Multiplikatoren (definiert für reguläres \bar{x})

$$\Lambda(\bar{x}) := \{\lambda \in \mathbb{R}^m : \text{ mit } \mathcal{L}_x(\bar{x}, 1, \lambda) = 0, \lambda(-y) \geq 0 \quad \forall y \in K(g(\bar{x}))\} \quad (3.9)$$

ist konvex und abgeschlossen.

Satz 3.22

$\bar{x} \in S$ ist regulär $\Leftrightarrow \Lambda(\bar{x}) \neq \emptyset$ und $\Lambda(\bar{x})$ ist beschränkt.

Beweis: Übungsaufgabe. □

Ist $\bar{x} \in S$ also regulär, dann ist $\Lambda(\bar{x})$ eine nicht-leere, konvexe und kompakte Teilmenge des \mathbb{R}^m .

Für den Einsatz von numerischen Verfahren zur Lösung der KKT-Bedingungen ist es wünschenswert, dass $\Lambda(\bar{x})$ einpunktig ist. Eine hinreichende Bedingung dafür erfordert eine verschärfte Regularität.

Definition 3.23

Ein Punkt $\bar{x} \in S$ heißt normal, wenn gilt

$$\text{Im}(g'(\bar{x})) - V = \mathbb{R}^m, V := K(g(\bar{x})) \cap (-K(g(\bar{x}))) \quad (3.10)$$

V ist der größte Untervektorraum, der in $K(g(\bar{x}))$ enthalten ist.

**Satz 3.24**

Ist $\bar{x} \in S$ normal und lokale Minimalstelle von $\min_{x \in S} f(x)$, dann hat $\Lambda(\bar{x})$ genau ein Element.

Beweis: Übungsaufgabe. □

Normalität ist für den Standardkegel $K = \mathbb{R}_+^k \times \{0\}_{m-k}$ äquivalent zur *Linear Independence Constraint Qualification (LICQ)*, siehe Paragraph 4.2 !

§ 3.2 Hinreichende Optimalitätsbedingungen

Betrachte wieder das Optimierungsproblem (3.4) in seiner allgemeinen Form.

Satz 3.25 (hinreichende Bed. 1. Ordnung)

Sei $\bar{x} \in S = \{x \in \mathbb{R}^n \mid g(x) \in K\}$ und sei $K(g(\bar{x}))$ abgeschlossen. Es gelte $f'(\bar{x})v > 0 \forall v \in L(S, \bar{x}) = \{v \in \mathbb{R}^n \mid g'(\bar{x})v \in K(g(\bar{x}))\}$. Dann $\exists c > 0$ und $\alpha > 0$ mit $f(x) \geq f(\bar{x}) + c\|x - \bar{x}\|$ für $\|x - \bar{x}\| \leq \alpha, x \in S$.

Beweis: (indirekt)

Annahme: \exists Folge $\{x_i\} \subset S, \lim_{i \rightarrow \infty} x_i = \bar{x}$ mit $f(x_i) < f(\bar{x}) + \frac{1}{i}\|x_i - \bar{x}\|$.

Wegen der Kompaktheit der Einheitskugel im \mathbb{R}^n existiere der Limes

$$v := \lim_{i \rightarrow \infty} \frac{x_i - \bar{x}}{\|x_i - \bar{x}\|} \in T(S, \bar{x}) \subset L(S, \bar{x}),$$

da $K(g(\bar{x}))$ abgeschlossen.

Außerdem gilt: $f'(\bar{x})v = \lim_{i \rightarrow \infty} \frac{f(x_i) - f(\bar{x})}{\|x_i - \bar{x}\|} \leq 0$, wegen $\frac{f(x_i) - f(\bar{x})}{\|x_i - \bar{x}\|} < \frac{1}{i} \forall i$.

⚡ Widerspruch zu $f'(\bar{x})v > 0, \forall v \in L(S, \bar{x})$. □

Betrachte nun den Kegel, auf dem diese hinreichenden Bed. 1. Ordnung nicht gelten:

$$C := \{v \in \mathbb{R}^n \mid f'(\bar{x})v \leq 0, g'(\bar{x})v \in K(g(\bar{x}))\} = \{v \in \mathbb{R}^n \mid f'(\bar{x})v \leq 0\} \cap L(S, \bar{x})$$

Voraussetzung: $f : \mathbb{R}^n \rightarrow \mathbb{R}, g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ 2-mal stetig differenzierbar in $\bar{x} \in S$.

Satz 3.26 (Ben-Tal, Lempio, Maurer, Zowe, 1979, 1981)

Sei $\bar{x} \in S$ und $K(g(\bar{x}))$ abgeschlossen. Hinreichende Bedingung 2. Ordnung (**SSC**, **Second-Order-Sufficient Condition**): Zu jedem $v \in C \setminus \{0\}$ gebe es ein $(\lambda_0, \lambda) \in \mathbb{R} \times \mathbb{R}^m$ mit

- (1) $\lambda_0 \geq 0, \lambda(-y) \geq 0 \forall y \in K(g(\bar{x}))$
- (2) $\lambda_0 f'(\bar{x}) + \lambda g'(\bar{x}) = 0$
- (3) $v^T(\lambda_0 f''(\bar{x}) + \lambda g''(\bar{x}))v > 0$

[(1) und (2): KKT-Bedingungen !]. Dann \exists Zahlen $c > 0, \alpha > 0$ mit

$$f(x) \geq f(\bar{x}) + c\|x - \bar{x}\|^2 \quad \forall \|x - \bar{x}\| \leq \alpha, x \in S \quad (3.11)$$

Beweis: (i.w. endlich-dimensional)

Angenommen Abschätzung (3.11) sei falsch. Dann \exists Folge $\{x_i\} \subset S$ mit $x_i \neq \bar{x}, \lim_{i \rightarrow \infty} x_i = \bar{x}$ und

$$f(x_i) < f(\bar{x}) + \frac{1}{i}\|x_i - \bar{x}\|^2 \quad (3.12)$$

Setze $v_i := x_i - \bar{x}$. Da $\{v \in \mathbb{R}^n : \|v\| = 1\}$ kompakt ist, so gelte o.E. $\lim_{i \rightarrow \infty} \frac{v_i}{\|x_i - \bar{x}\|} =: v$.

Mit (3.12) folgt $f'(\bar{x})v = \lim_{i \rightarrow \infty} \underbrace{\frac{f(x_i) - f(\bar{x})}{\|x_i - \bar{x}\|}}_{< \frac{1}{i}\|x_i - \bar{x}\|} \leq 0$.

Wegen der Abgeschlossenheit von $K(g(\bar{x}))$ gilt

$$g'(\bar{x})v = \lim_{i \rightarrow \infty} \underbrace{\frac{g(x_i) - g(\bar{x})}{\|x_i - \bar{x}\|}}_{\in K(g(\bar{x}))} \in K(g(\bar{x})),$$

also folgt $v \in C$. Taylorentwicklung 2. Ordnung ($v_i = x_i - \bar{x}$):

$$f'(\bar{x})v_i + \frac{1}{2}v_i^\top f''(\bar{x})v_i + o(\|v_i\|^2) = f(x_i) - f(\bar{x}) \leq \frac{1}{i}\|x_i - \bar{x}\|^2 \quad (\text{nach (3.12)}) \quad (3.13)$$

$$g'(\bar{x})v_i + \frac{1}{2}v_i^\top g''(\bar{x})v_i + o(\|v_i\|^2) = g(x_i) - g(\bar{x}) \in K(g(\bar{x})) \quad (3.14)$$

Multipliziere (3.13) mit $\lambda_0 \geq 0$ und (3.14) mit λ . Eine Addition von (3.13) und (3.14) ergibt dann wegen $\lambda \underbrace{(g(x_i) - g(\bar{x}))}_{\in K(g(\bar{x}))} \leq 0$ (wg. Bed. (1)) und mit Bed. (1) und

(2) des Satzes:

$$\begin{aligned} \frac{1}{2}v_i^\top (\lambda_0 f''(\bar{x}) + \lambda g''(\bar{x}))v_i &\leq \frac{\lambda_0}{i}\|x_i - \bar{x}\|^2 \\ \Rightarrow \frac{1}{2} \frac{v_i^\top}{\|x_i - \bar{x}\|} (\lambda_0 f''(\bar{x}) + \lambda g''(\bar{x})) \frac{v_i}{\|x_i - \bar{x}\|} &\leq \frac{\lambda_0}{i} \end{aligned}$$

für $i \rightarrow \infty$ folgt $\frac{1}{2}v^\top (\lambda_0 f''(\bar{x}) + \lambda g''(\bar{x}))v \leq 0$ im Widerspruch zur Bedingung (3) des Satzes. \square

Berechnung des Kegels C :

Sei der Satz von KKT für genau ein $\lambda \in \mathbb{R}^m$ erfüllt an der Stelle \bar{x} , \bar{x} regulär ($\rightarrow \lambda_0 > 0$, teile in KKT-Bed. durch λ_0).

$$(KKT) \quad f'(\bar{x}) + \lambda g'(\bar{x}) = 0, \quad \lambda(-y) \geq 0 \quad \forall y \in K(g(\bar{x}))$$

Für $v \in C$ gilt: $f'(\bar{x})v \leq 0, \quad g'(\bar{x})v \in K(g(\bar{x}))$

$$\begin{aligned} \text{Mit KKT folgt: } 0 \geq f'(\bar{x})v &= -\lambda g'(\bar{x})v = \lambda(-g'(\bar{x})v) \geq 0 \\ &\Rightarrow \lambda g'(\bar{x})v = 0 = f'(\bar{x})v \end{aligned}$$

Dann hat C die Darstellung:

$$C = \{v \in \mathbb{R}^n \mid \lambda g'(\bar{x})v = 0, \quad g'(\bar{x})v \in K(g(\bar{x}))\}$$

Speziell für Gleichungen, d.h. $g(x) = 0$ ($K = \{0\}$):

$$C = \{v \in \mathbb{R}^n \mid g'(\bar{x})v = 0\}$$

Standardproblem

$$\begin{aligned} \min_x \quad & f(x) \\ \text{unter} \quad & g_i(x) \leq 0, \quad i = 1, \dots, k \\ & g_i(x) = 0, \quad i = k+1, \dots, m \end{aligned}$$

$I(\bar{x})$: aktive Indizes (aktive Ungleichungen)

$$J(\bar{x}) := I(\bar{x}) \cup \{k+1, \dots, m\}$$

Für $\lambda = (\lambda_1, \dots, \lambda_m)$ gilt wegen der Komplementaritätsbedingung (siehe Bem. 3.21) $\lambda_i \geq 0$ für $i \in I(\bar{x})$ und $\lambda_i = 0$ für $i \notin J(\bar{x})$.

Wegen $K(g(\bar{x})) = \{y \in \mathbb{R}^m \mid y_i \leq 0, \quad i \in I(\bar{x}); \quad y_i = 0, \quad i = k+1, \dots, m\}$ hat man für den Kegel C die Darstellung

$$\begin{aligned} C = \{v \in \mathbb{R}^n \mid & g'_i(\bar{x})v \leq 0 \quad \forall i \in I(\bar{x}) \text{ mit } \lambda_i = 0; \\ & g'_i(\bar{x})v = 0 \quad \forall i \in I(\bar{x}) \text{ mit } \lambda_i > 0; \\ & g'_i(\bar{x})v = 0 \quad i = k+1, \dots, m\}, \end{aligned}$$

da $\lambda g'(\bar{x})v = 0$ komponentenweise gelten muss, denn alle Summanden in $\sum_{i=1}^m \lambda_i g'_i(\bar{x})v = 0$ haben wegen $\lambda(-y) \geq 0$ gleiches Vorzeichen.

Wie wir im folgenden Abschnitt zur Sensitivitätsanalyse sehen werden, fordert man zur expliziten Berechnung von C in der Praxis i.d.R. zusätzlich zu den SSC in Satz 3.26 die sog. strikte Komplementarität ($\lambda_i > 0 \quad \forall i \in I(\bar{x})$) und kann damit den Kegel C über den Kern der Matrix der Gradienten $g'_i(\bar{x}), i \in J(\bar{x})$ der Gleichungs- und aktiven Ungleichungsbeschränkungen beschreiben.

§ 3.3 Sensitivitätsanalyse

Sei $f : \mathbb{R}^n \times \mathbb{R}^q \longrightarrow \mathbb{R}$ 2-mal stetig differenzierbar in einer Umgebung von $(\bar{x}, p_0) \in \mathbb{R}^n \times \mathbb{R}^q$. Notation: $f_x(\bar{x}, p_0)$, $f_{xx}(\bar{x}, p_0)$ partielle Ableitungen nach x (Gradient und Hesse-Matrix).

Satz 3.27 (Sensitivitätssatz, unbeschränkter Fall $\min_x f(x, p_0)$)

Der Punkt \bar{x} genüge den hinreichenden Optimalitätsbedingungen 2. Ordnung für das parametrische Optimierungsproblem

$$P(p_0) : \min_{x \in \mathbb{R}^n} f(x, p_0),$$

d.h. $f_x(\bar{x}, p_0) = 0$ und $f_{xx}(\bar{x}, p_0) > 0$ (pos. def.).

Dann gibt es eine Umgebung P_0 von p_0 und eine stetig differenzierbare Funktion $x : P_0 \longrightarrow \mathbb{R}^n$ mit

- (1) $x(p_0) = \bar{x}$
- (2) $x(p)$ strenge lokale Minimalstelle von $P(p) : \min_{x \in \mathbb{R}^n} f(x, p)$
- (3) $\frac{d}{dp} f(x(p), p)|_{p=p_0} = f_p(\bar{x}, p_0) = \frac{\partial f}{\partial p}(\bar{x}, p_0)$
- (4) $\frac{dx}{dp}(p_0) = -f_{xx}(\bar{x}, p_0)^{-1} f_{xp}(\bar{x}, p_0)$

Beweis: Setze $F(x, p) := f(x, p) : \mathbb{R}^n \times \mathbb{R}^q \longrightarrow \mathbb{R}$.

Wegen $F(\bar{x}, p_0) = f(\bar{x}, p_0) = 0$ und $F_x(\bar{x}, p_0) = \frac{\partial F}{\partial x}(\bar{x}, p_0) = f_{xx}(\bar{x}, p_0) > 0$, also regulär, sind die Voraussetzungen des Satzes über implizite Funktionen erfüllt, d.h. \exists Umgebung P_0 von $p_0 \in \mathbb{R}^k$ und eine C^1 -Abbildung $x : P_0 \longrightarrow \mathbb{R}^n$ mit $F(x(p), p) \equiv 0$ (d.h. $F(x(p), p) = 0 \forall p \in P_0$). Für P_0 genügend klein bleibt aufgrund der Stetigkeit $f_{xx}(x(p), p) > 0 \implies x(p)$ strenge lokale Minimalstelle von $P(p) \Rightarrow$ Aussage (1) und (2).

Zu (3): $\frac{d}{dp} f(x(p), p)|_{p=p_0} = \underbrace{f_x(\bar{x}, p_0)}_{=0} \frac{\partial x}{\partial p}(p_0) + f_p(\bar{x}, p_0) = f_p(\bar{x}, p_0)$.

Zu (4): Differenziere die Gleichung $F(x(p), p) = 0 \forall p \in P_0$:

$0 = \frac{d}{dp} F(x(p), p) = F_x(\bar{x}, p_0) \frac{\partial x}{\partial p}(p_0) + F_p(\bar{x}, p_0)$, Auflösen nach $\frac{\partial x}{\partial p}(p_0)$ ergibt (4). \square

Anwendung in der „Echtzeit“-Optimierung:

$$\begin{aligned} x(p) &\approx x(p_0) + \frac{dx}{dp}(p_0)(p - p_0) \quad [\text{online-Berechnung}] \\ x(p_0) &= \bar{x}, \quad \frac{dx}{dp}(p_0) \quad [\text{offline-Berechnung}] \end{aligned}$$

Beschränkter Fall: Standardproblem

$$\begin{aligned} \min \quad & f(x) \\ \text{unter} \quad & g_i(x) \leq 0, \quad i = 1, \dots, k \\ & g_i(x) = 0, \quad i = k + 1, \dots, m \end{aligned}$$

Annahme: Die Daten des Problems $f(x)$, $g_i(x)$ unterliegen Störungen in Form von Parametern. Spezialfall: Lineare Störung

$$\begin{aligned} \min \quad & f(x) \\ \text{unter} \quad & g_i(x) \leq p_i, \quad i = 1, \dots, k \\ & g_i(x) = p_i, \quad i = k + 1, \dots, m \end{aligned}$$

$p = (p_1, \dots, p_k, \dots, p_m) \in \mathbb{R}^m$ Störungsparameter; $p = 0$ ungestörtes Problem.

Betrachte allgemein die parametrische Schar von Optimierungsproblemen

$$\begin{aligned} & \min_x f(x, p) \\ & \text{unter } g_i(x, p) \leq 0, \quad i = 1, \dots, k \\ & \quad \quad g_i(x, p) = 0, \quad i = k + 1, \dots, m \\ & p \in \mathbb{R}^q; \quad f, g_i : \mathbb{R}^n \times \mathbb{R}^q \longrightarrow \mathbb{R} \end{aligned}$$

oder allgemeines Störungsproblem

$$\min \{f(x, p) : g(x, p) \in K\}; \quad K \subset \mathbb{R}^m \text{ konvex}$$

Zulässige Menge $S(p) = \{x \in \mathbb{R}^n : g(x, p) \in K\}$

Optimalwertfunktion $v(p) := \inf \{f(x, p) : g(x, p) \in K\}, v : \mathbb{R}^q \rightarrow \mathbb{R} \cup \{-\infty, +\infty\}$

Ungestörtes Problem: wähle $p = p_0 \in \mathbb{R}^q$ (nomineller Parameter)

Sensitivitätsanalyse: 2 Problemstellungen

- (1) Bette Lsg. \bar{x} des nominellen Problems ein in eine C^1 -Familie $x(p)$ von Lösungen der gestörten Probleme
- (2) Berechne Abschätzungen für die Richtungsableitungen von $x(p)$ nach dem Parameter p

Sensitivität der optimalen Lösung im beschränkten Standardfall

$$\begin{aligned} P(p) \quad & \min_x \{f(x, p) : g_i(x, p) \leq 0, \quad i = 1, \dots, k \\ & \quad \quad g_i(x, p) = 0, \quad i = k + 1, \dots, m\} \end{aligned}$$

Nominelles Problem: $P(p_0)$

Notation: $I(\bar{x}) = \{i \in \{1, \dots, k\} : g_i(\bar{x}, p_0) = 0\}$

$$J(\bar{x}) = \{k + 1, \dots, m\} \cup I(\bar{x})$$

$$k_0 := \#I(\bar{x}) \text{ (Kardinalität)}$$

$$m_0 := \#J(\bar{x}) = k_0 + m - k$$

Satz 3.28 (Sensitivitätssatz, beschränkter Fall)

Sei $\bar{x} \in S(p_0)$, $p_0 \in \mathbb{R}^q$ eine lokale Minimalstelle des nominellen Problems $P(p_0)$ und sei \bar{x} normal (d.h. die Gradienten $g_{ix}(\bar{x}, p_0)$, $i \in J(\bar{x})$ sind linear unabhängig, s. Übungsaufgabe). Es gebe Multiplikatoren $\bar{\lambda}_i$, $i \in J(\bar{x})$, so dass die hinreichenden Bedingungen 2. Ordnung (SSC) aus Satz 3.26 gelten mit

- a) $\bar{\lambda}_i > 0$, $i \in I(\bar{x})$ (strikte Komplementarität)
- b) $f_x(\bar{x}, p_0) + \sum_{i \in J(\bar{x})} \bar{\lambda}_i g_{ix}(\bar{x}, p_0) = 0$
- c) $v^T (f_{xx}(\bar{x}, p_0) + \sum_{i \in J(\bar{x})} \bar{\lambda}_i g_{ixx}(\bar{x}, p_0)) v > 0 \quad \forall v \neq 0, v \in C = \{v \in \mathbb{R}^n : g_{ix}(\bar{x}, p_0)v = 0, i \in J(\bar{x})\}$

Dann \exists Umgebung $P_0 \subset \mathbb{R}^q$ von p_0 und stetig differenzierbare Abbildungen

$$x : P_0 \longrightarrow \mathbb{R}^n$$

$$\lambda_i : P_0 \longrightarrow \mathbb{R}, \quad i \in J(\bar{x})$$

mit folgenden Eigenschaften:

- (1) $x(p_0) = \bar{x}$, $\lambda_i(p_0) = \bar{\lambda}_i \quad (i \in J(\bar{x}))$

- (2) Die Punkte $x(p)$, $\lambda_i(p)$, $i \in J(\bar{x})$ erfüllen die SSC für das gestörte Problem $P(p) \forall p \in P_0$.

Beweis: Vom Konzept analog zum unbeschränkten Fall.
Definiere die Funktion der ‘aktiven Gleichungen’

$$G : \mathbb{R}^n \times \mathbb{R}^q \rightarrow \mathbb{R}^{m_0}, \quad G(x, p) = \left((g_i(x, p))_{i \in I(\bar{x})}, g_{k+1}(x, p), \dots, g_m(x, p) \right)^T$$

$\bar{x} \in S(p_0)$ ist n.V. lokale Minimalstelle des Problems

$$\min_{x \in \mathbb{R}^n} \{f(x, p) \mid G(x, p_0) = 0\},$$

da $\lambda_i = 0, i \notin J(\bar{x})$. Die gesuchte Minimalstelle $x = x(p)$ soll stetig differenzierbar sein. Also muss gelten

$$g_i(x(p), p) < 0 \text{ für } i \notin J(\bar{x}), \|p - p_0\| \text{ genügend klein.}$$

Daher ist $x = x(p)$ lokale Minimalstelle eines gestörten Problems mit Gleichungen

$$\min_{x \in \mathbb{R}^n} \{f(x, p) \mid G(x, p) = 0\} \quad (3.15)$$

Die zugehörige Lagrange-Funktion ist

$$L(x, \nu, p) = f(x, p) + \nu G(x, p), \quad \nu \in \mathbb{R}^{m_0} \quad (3.16)$$

und liefert für die KKT-Bed. ein Gleichungssystem in $x = x(p)$ und $\nu(p) = (\lambda_i(p))_{i \in J(\bar{x})}$:

$$F(x, \nu^\top, p) := \begin{pmatrix} L_x(x, \nu, p) \\ G(x, p) \end{pmatrix} = 0 \quad (3.17)$$

Hierbei ist $F : \mathbb{R}^{n+m_0} \times \mathbb{R}^q \rightarrow \mathbb{R}^n \times \mathbb{R}^{m_0}$.

Für $p = p_0$ gilt $F(\bar{x}, \bar{\nu}^\top, p_0) = 0$, $\bar{\nu} = (\bar{\lambda}_i^\top)_{i \in J(\bar{x})}$. Wir zeigen: F ist stetig differenzierbar in einer Umgebung von $(\bar{x}, \bar{\nu}^\top, p_0)$.

Partielle JACOBI-Matrix von F bzgl. (x, ν^\top) im Punkte $(\bar{x}, \bar{\nu}^\top, p_0)$:

$$A_0 := \frac{\partial F}{\partial (x, \nu^\top)}(\bar{x}, \bar{\nu}^\top, p_0) = \begin{pmatrix} L_{xx}(\bar{x}, \bar{\nu}, p_0) & G_x^\top(\bar{x}, p_0) \\ \underbrace{G_x(\bar{x}, p_0)}_{\text{diff nach } x} & \underbrace{0}_{\text{diff nach } \nu^\top} \end{pmatrix} \in M(n + m_0, n + m_0) \quad (3.18)$$

Zeige: A_0 ist regulär, falls die SSC a-c) in \bar{x} gelten. Dazu sei $(v, w) \in \mathbb{R}^n \times \mathbb{R}^{m_0}$ mit

$$A_0 \begin{pmatrix} v \\ w \end{pmatrix} = \begin{pmatrix} L_{xx}(\bar{x}, \bar{\nu}, p_0)v + G_x^\top(\bar{x}, p_0)w \\ G_x(\bar{x}, p_0)v \end{pmatrix} = 0 \quad (3.19)$$

und zeige, dass das nur für $(v, w) = (0, 0)$ möglich ist. Multipliziere (3.19) mit $(v, 0)^\top$ von links

$$v^\top L_{xx}(\bar{x}, \bar{\nu}, p_0)v + \underbrace{v^\top G_x^\top(\bar{x}, p_0)}_{=(G_x(\bar{x}, p_0)v)^\top = 0} w = v^\top L_{xx}(\bar{x}, \bar{\nu}, p_0)v = 0 \quad (3.20)$$

wegen $G_x(\bar{x}, p_0)v = 0$ (2. Gleichung in (3.19)) folgt $v \in C$. Falls $v \neq 0 \notin$ ist das ein Widerspruch zu den SSC (Bed. c), also muss gelten $v = 0$, dann folgt aber, dass A_0 regulär ist, da mit $v = 0$ aus der 1. Gleichung in (3.19) folgt $G_x^\top(\bar{x}, p_0)w = 0 \Rightarrow w = 0$, denn $G_x^\top(\bar{x}, p_0)$ hat nach Voraussetzung (Normalität von \bar{x} !) maximalen Rang m_0 , da $g_{ix}(\bar{x}, p_0)$ linear unabhängig.

Anwendung des Satzes über implizite Funktionen:

\exists Umgebung $P_0 \subset \mathbb{R}^q$ von p_0 und C^1 -Abb. $x : P_0 \rightarrow \mathbb{R}^n$, $\lambda_i : P_0 \rightarrow \mathbb{R}$, $i \in J(\bar{x})$ mit

$$F(x(p), \nu(p)^\top, p) = 0 \quad \forall p \in P_0 \quad (3.21)$$

wobei $\nu(p) = (\lambda_i(p))_{i \in J(\bar{x})}$. Es gilt: $x(p_0) = \bar{x}$, $\lambda_i(p_0) = \bar{\lambda}_i$, $i \in J(\bar{x})$.

Zeige: Die Vektoren $x(p)$, $(\lambda_i(p))_{i \in J(\bar{x})}$ erfüllen die SSC für das gestörte Problem für P_0 geeignet "klein".

Begründung: $\lambda_i(p_0) = \bar{\lambda}_i > 0 \forall i \in I(\bar{x})$ und $g_i(\bar{x}, p_0) < 0 \forall i \notin J(\bar{x}) \Rightarrow \lambda_i(p) > 0 \forall i \in I(\bar{x})$, $g_i(x(p), p) < 0 \forall i \notin J(\bar{x})$ für $\|p - p_0\|$ genügend klein (Stetigkeitsargument). Also ist $x(p) \in S(p)$ (zulässig). Außerdem gilt wg. (3.21) $L_x(x(p), \nu(p), p) = 0 \forall p \in P_0$.

Überlege (Stetigkeitsargument):

$$v^\top L_{xx}(x(p), \nu(p), p) v > 0 \quad \forall v \neq 0, v \in C$$

$$C = \{v \in \mathbb{R}^n \mid G_x(x(p), p) v = 0\}.$$

□

Korollar 3.29 (zum Sensitivitätssatz)

$G : \mathbb{R}^n \times \mathbb{R}^q \rightarrow \mathbb{R}^{m_0}$, $G(x, p) := (g_i(x, p))_{i \in J(\bar{x})}$

- i) Mit der $(n + m_0) \times (n + m_0)$ -Matrix $A_0 := \begin{pmatrix} L_{xx}(\bar{x}, \bar{\lambda}, p_0) & G_x(\bar{x}, p_0)^\top \\ G_x(\bar{x}, p_0) & 0 \end{pmatrix}$
und der $(n + m_0) \times q$ -Matrix $B_0 := \begin{pmatrix} L_{xp}(\bar{x}, \bar{\lambda}, p_0) \\ G_p(\bar{x}, p_0) \end{pmatrix}$

berechnet man $x'(p_0) = \frac{dx}{dp}(p_0)$ und $\lambda'(p_0) = \frac{d\lambda}{dp}(p_0)$ gemäß der Formel

$$\begin{pmatrix} x'(p_0) \\ \lambda'(p_0) \end{pmatrix} = -A_0^{-1} B_0$$

- ii) Verallgemeinerte Schattenpreisformel

$$\frac{d}{dp} f(x(p), p)|_{p=p_0} = L_p(\bar{x}, \bar{\lambda}, p_0) = f_p(\bar{x}, p_0) + \bar{\lambda} G_p(\bar{x}, p_0)$$

Für den Spezialfall

$$P(p) : \min_x \quad \{f(x) \mid g_i(x) \leq p_i, \quad i = 1, \dots, k \\ g_i(x) = p_i, \quad i = k + 1, \dots, m\}$$

gilt $\frac{d}{dp_i} f(x(p))|_{p=p_0} = L_{p_i}(\bar{x}, \bar{\lambda}, p_0) = \begin{cases} -\bar{\lambda}_i & i \in J(\bar{x}) \\ 0 & i \notin J(\bar{x}) \end{cases}$ (Analogie zu LP ! Damit folgt

Formel (5.7))

Beweis: Analog zum unbeschränkten Fall.

zu i): Differentiation von (3.21) ergibt:

$$0 = \frac{d}{dp} F(x(p), \nu^\top(p), p) \Big|_{p=p_0} = \frac{\partial F}{\partial(x, \nu^\top)}(\bar{x}, \bar{\nu}^\top, p_0) \begin{pmatrix} x'(p_0) \\ \nu'(p_0)^\top \end{pmatrix} + \frac{\partial F}{\partial p}(\bar{x}, \bar{\nu}^\top, p_0) = A_0 \begin{pmatrix} x'(p_0) \\ \nu'(p_0)^\top \end{pmatrix} + B_0 \text{ (LGS zur Bestimmung von } (x'(p_0), \nu'(p_0)) \text{)}$$

\Rightarrow i), da A_0 regulär (Satz über implizite Funktion). Aus numerischen Gründen wird eine Invertierung von A_0 vermieden und das LGS für $\begin{pmatrix} x'(p_0) \\ \nu'(p_0)^\top \end{pmatrix}$ gelöst.

zu ii): Zunächst gilt $G(x(p), p) = 0 \forall p \in P_0$.

$$0 = \frac{d}{dp} G(x(p), p) \Big|_{p=p_0} = G_x(\bar{x}, p_0) x'(p_0) + G_p(\bar{x}, p_0)$$

Zudem gilt nach KKT: $f_x(\bar{x}, p_0) = -\bar{\nu} G_x(\bar{x}, p_0)$

Damit folgt: $\frac{d}{dp} f(x(p), p) \Big|_{p=p_0} = f_x(\bar{x}, p_0) x'(p_0) + f_p(\bar{x}, p_0) = -\nu \underbrace{G_x(\bar{x}, p_0) x'(p_0)}_{=-G_p(\bar{x}, p_0)} + f_p(\bar{x}, p_0) = \nu G_p(\bar{x}, p_0) + f_p(\bar{x}, p_0) = L_p(\bar{x}, \bar{\nu}, p_0)$ \square

Anwendung in der Echtzeit-Optimierung: Berechne

$$\begin{pmatrix} x(p) \\ \nu(p)^\top \end{pmatrix} \approx \begin{pmatrix} x(p_0) = \bar{x} \\ \nu(p_0)^\top = \bar{\nu}^\top \end{pmatrix} + \begin{pmatrix} x'(p_0) \\ \nu'(p_0)^\top \end{pmatrix} (p - p_0)$$

für "leicht" gestörtes Optimierungsproblem aus den "Daten" des nominellen Problems.

§ 4 Standardproblem der Nichtlinearen Optimierung im \mathbb{R}^n

Wir betrachten in diesem Abschnitt das allgemeine nichtlineare Optimierungsproblem (vgl. (2.2)), auch NLP (Nonlinear Program) genannt:

$$\left. \begin{array}{ll} \text{Minimiere} & f(x) \quad \text{über } x \in \mathbb{R}^n \\ \text{sodass} & g_i(x) \leq 0, \quad i = 1, \dots, k \\ \text{und} & h_j(x) = 0, \quad j = 1, \dots, m-k. \end{array} \right\} \quad (4.1)$$

mit $0 \leq k \leq m$.

Wir nehmen im folgenden an, dass f, g und h stetig differenzierbar auf \mathbb{R}^n sind. Zur Erinnerung:

$$S = \{x \in \mathbb{R}^n : g_i(x) \leq 0 \text{ für alle } i = 1, \dots, k, \quad h_j(x) = 0 \text{ für alle } j = 1, \dots, m-k\} \quad (4.2)$$

ist die **zulässige Menge** von (4.1).

§ 4.1 Notwendige Optimalitätsbedingungen der nichtlinearen Optimierung im \mathbb{R}^n

Definition 4.1 (Lagrange-Funktion für (4.1))

Die Funktion

$$\begin{aligned} \mathcal{L}(x, \lambda, \mu) &:= f(x) + \sum_{i=1}^k \mu_i g_i(x) + \sum_{j=1}^{m-k} \lambda_j h_j(x) \\ &= f(x) + \mu^\top g(x) + \lambda^\top h(x) \end{aligned}$$

heißt die zu (4.1) gehörige **Lagrange-Funktion**. \diamond

Definition 4.2 (KKT-Bedingungen und -Punkt)

(a) Die Bedingungen

$$\mathcal{L}_x(x, \lambda, \mu) = 0 \quad (4.3a)$$

$$h(x) = 0 \quad (4.3b)$$

$$\mu \geq 0, \quad g(x) \leq 0, \quad \mu^\top g(x) = 0 \quad (4.3c)$$

heißen die **Karush-Kuhn-Tucker-Bedingungen** (kurz: **KKT-Bedingungen**) des Problems (4.1). Dabei ist

$$\begin{aligned} \mathcal{L}_x(x, \lambda, \mu) &= \nabla f(x) + \sum_{i=1}^k \mu_i \nabla g_i(x) + \sum_{j=1}^{m-k} \lambda_j \nabla h_j(x) \\ &= \nabla f(x) + \underbrace{g'(x)^\top}_{\text{JACOBI-Matrix}} \mu + h'(x)^\top \lambda \end{aligned}$$

der Gradient von $\mathcal{L}(x, \lambda, \mu)$ bzgl. x .

(b) Ein Punkt $(x^*, \lambda^*, \mu^*) \in \mathbb{R}^n \times \mathbb{R}^{m-k} \times \mathbb{R}^k$, der (4.3) erfüllt, heißt ein **KKT-Punkt** von (4.1). Die Vektoren λ^* und μ^* heißen dann **Lagrange-Multiplikatoren** zu den Beschränkungen $h(x^*) = 0$ bzw. $g(x^*) \leq 0$. \diamond

Bemerkung 4.3

Sei (x^*, λ^*, μ^*) ein KKT-Punkt.

(a) Die **Komplementaritätsbedingung** (4.3c) ist äquivalent zu

$$\mu_i^* \geq 0, \quad g_i(x^*) \leq 0, \quad \mu_i^* g_i(x^*) = 0 \quad \text{für alle } i = 1, \dots, k,$$

d.h., mindestens eine der Zahlen μ_i^* und $g_i(x^*)$ ist gleich null. Insbesondere gehört also zu einer inaktiven Beschränkung $g_i(x^*) < 0$ ein Lagrange-Multiplikator $\mu_i^* = 0$.

(b) Die Bedingung (4.3a) besagt gerade, dass

$$-\nabla f(x^*) = \sum_{i=1}^k \mu_i^* \nabla g_i(x^*) + \sum_{j=1}^{m-k} \lambda_j^* \nabla h_j(x^*)$$

gilt mit Skalaren $\mu_i^* \geq 0$ und $\mu_i^* = 0$ für i mit $g_i(x^*) < 0$ wegen (4.3c). Außerdem sichern (4.3b) und (4.3c) die Zulässigkeit $g(x^*) \leq 0$ und $h(x^*) = 0$.

(c) Durch die KKT-Bedingungen wird die Optimierungsaufgabe (4.1) in ein nichtlineares Gleichungs- und Ungleichungssystem (4.3) transformiert, das als Basis für viele numerische Verfahren dient.

(d) Man nennt die KKT-Bedingungen auch **notwendige Bedingungen 1. Ordnung**, da sie (nur) erste Ableitungen von f , g und h verwenden.

§ 4.2 Notwendige Bedingungen unter LICQ**Definition 4.4 (LICQ)**

Sei x_0 zulässiger Punkt des Optimierungsproblems (4.1) und $I(x_0)$ die Menge der aktiven Indizes, d.h. die Indizes derjenigen Ungleichungen $g_i(x) \leq 0$, für die in x^* die

Gleichheit $g_i(x) = 0$ erfüllt ist. Man sagt: x_0 erfüllt die **Regularitätsbedingungen der linearen Unabhängigkeit** (kurz: **LICQ**, Linear Independence Constraint Qualification), wenn gilt: Die Gradienten $\{\nabla h_j(x_0)\}_{j=1,\dots,m-k} \cup \{\nabla g_i(x_0)\}_{i \in I(x_0)}$ sind linear unabhängig. \diamond

Satz 4.5 (KKT-Bedingungen unter LICQ)

Es sei x^* ein lokales Optimum von (4.1), das die LICQ erfüllt. Dann existieren *eindeutig bestimmte* Lagrange-Multiplikatoren $\lambda^* \in \mathbb{R}^{m-k}$ und $\mu^* \in \mathbb{R}^k$, so dass (x^*, λ^*, μ^*) ein KKT-Punkt von (4.1) ist.

§ 4.3 Notwendige Bedingungen unter MFCQ

Definition 4.6 (MFCQ)

Sei x_0 zulässiger Punkt des Optimierungsproblems (4.1) und $I(x_0)$ die Menge der aktiven Indizes. Man sagt: x_0 erfüllt die **Regularitätsbedingungen von Mangasarian-Fromowitz** (kurz: **MFCQ**), wenn gilt:

- (a) Die Gradienten $\{\nabla h_j(x_0)\}_{j=1}^{m-k}$ sind linear unabhängig.
- (b) Es existiert ein Vektor $d \in \mathbb{R}^n$ mit

$$\begin{aligned} \nabla g_i(x_0)^\top d &< 0 \quad \text{für alle } i \in I(x_0) \\ \nabla h_j(x_0)^\top d &= 0 \quad \text{für alle } j = 1, \dots, m-k, \end{aligned} \quad \diamond$$

Satz 4.7 (KKT-Bedingungen unter MFCQ)

Es sei x^* ein lokales Optimum von (4.1), das die MFCQ erfüllt. Dann existieren *nicht notwendig eindeutig bestimmte* Lagrange-Multiplikatoren $\lambda^* \in \mathbb{R}^{m-k}$ und $\mu^* \in \mathbb{R}^k$, sodass (x^*, λ^*, μ^*) ein KKT-Punkt von (4.1) ist.

§ 4.4 Notwendige Bedingungen für konvexe Optimierungsprobleme unter Slater-Bedingungen

Wir betrachten das Optimierungsproblem

$$\left. \begin{aligned} &\text{Minimiere} && f(x) && \text{über } x \in \mathbb{R}^n \\ &\text{sodass} && g_i(x) \leq 0, && i = 1, \dots, k \\ &\text{und} && Ax = b. \end{aligned} \right\} \quad (4.4)$$

Dabei sind

- (a) $f : \mathbb{R}^n \rightarrow \mathbb{R}$ stetig differenzierbar und konvex,
- (b) $g_i : \mathbb{R}^n \rightarrow \mathbb{R}$ für $i = 1, \dots, m$ stetig differenzierbar und konvex,
- (c) $A \in \mathbb{R}^{(m-k) \times n}$ und $b \in \mathbb{R}^{m-k}$.

Beachte: (4.4) ist eine konvexe Optimierungsaufgabe.

Definition 4.8 (Slater-Bedingung)

Die Aufgabe (4.4) erfüllt die **Regularitätsbedingung von Slater**, wenn ein $x_0 \in \mathbb{R}^n$ existiert mit

$$g_i(x_0) < 0 \quad \text{für } i = 1, \dots, k \quad \text{und} \quad Ax_0 = b. \quad (4.5)$$

[D.h., x_0 ist **strikt zulässig** bzgl. der Ungleichungen und zulässig bzgl. der Gleichungen.] Ein solches x_0 heißt **Slater-Punkt**. \diamond

Vorteil: x^* muss nicht bekannt sein.

Satz 4.9 (KKT-Bedingungen unter Slater-Bedingung)

Es sei x^* ein [lokales=globales] Optimum von (4.4), und es sei die Slater-Bedingung mit einem $x_0 \in \mathbb{R}^n$ erfüllt. Dann existieren Lagrange-Multiplikatoren $\lambda^* \in \mathbb{R}^{m-k}$ und $\mu^* \in \mathbb{R}^k$, sodass (x^*, λ^*, μ^*) ein KKT-Punkt von (4.4) ist, d.h.,

$$\nabla f(x^*) + g'(x^*)^\top \mu^* + A^\top \lambda^* = 0 \quad (4.6a)$$

$$Ax^* = b \quad (4.6b)$$

$$\mu^* \geq 0, \quad g(x^*) \leq 0, \quad (\mu^*)^\top g(x^*) = 0 \quad (4.6c)$$

Korollar 4.10 (KKT-Bedingungen unter linearisierter Slater-Bedingung)

Alternativ zur Slater-Bedingung (4.5) kann man in Satz 4.9 an der Minimalstelle x^* fordern: Es existiert $d \in \mathbb{R}^n$ mit

$$g_i(x^*) + \nabla g_i(x^*)^\top d < 0 \quad \text{und} \quad Ad = 0. \quad (4.7)$$

Diese Bedingung heißt die **linearisierte Regularitätsbedingung von Slater** (siehe auch (3.2) und (3.3)).

Des weiteren zeigt folgender Satz, dass die KKT-Bedingungen für das konvexe Optimierungsproblem (4.4) bereits hinreichend sind.

Satz 4.11 (hinreichende Bedingung)

Sei $(x^*, \lambda^*, \mu^*) \in \mathbb{R}^n \times \mathbb{R}^{m-k} \times \mathbb{R}^k$ ein KKT-Punkt des konvexen Optimierungsproblems (4.4). Dann ist x^* ein (globales=lokales) Minimum von (4.4).

Beweis: Sei $x \in \mathbb{R}^n$ eine zulässiger Punkt für das konvexe Optimierungsproblem (4.4) und $(x^*, \lambda^*, \mu^*) \in \mathbb{R}^n \times \mathbb{R}^{m-k} \times \mathbb{R}^k$ ein KKT-Punkt von (4.4). Dann folgt mit Hilfe der KKT-Bedingungen, der Konvexität von f und Satz 2.24:

$$\begin{aligned} f(x) &\geq f(x^*) + \nabla f(x^*)^\top (x - x^*) \\ &= f(x^*) - \sum_{i=1}^k \mu_i^* \nabla g_i(x^*)^\top (x - x^*) - (\lambda^*)^\top A(x - x^*) \\ &= f(x^*) - \sum_{i \in I(x^*)} \mu_i^* \nabla g_i(x^*)^\top (x - x^*) - (\lambda^*)^\top (b - b) \\ &\geq f(x^*) - \sum_{i \in I(x^*)} \mu_i^* (g_i(x) - g_i(x^*)) = f(x^*) - \sum_{i \in I(x^*)} \mu_i^* g_i(x) \\ &\geq f(x^*). \end{aligned}$$

Somit ist x^* ein globales Minimum von (4.4). □

§ 5 Lineare Optimierung als Spezialfall

§ 5.1 Grundbegriffe und Konzepte

Beispiel 5.1 (Schuhfabrik)

Eine Schuhfabrik stellt Herren- und Damenschuhe her

	Damen	Herren	verfügbare Aktivität
Herstellungszeit	20	10	8000
Maschinenbedarf	4	5	2000
Lederbedarf	6	15	4500
Gewinn pro Einheit	16	32	

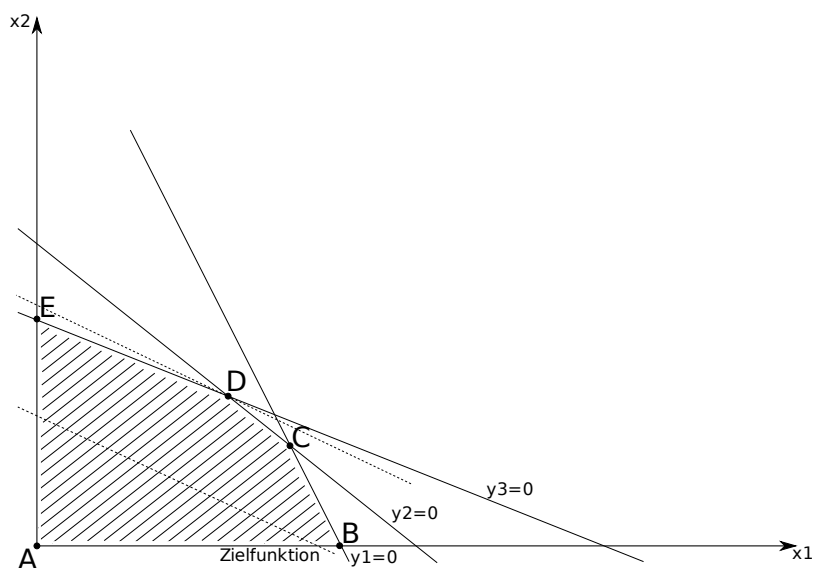
Zugehöriges lineares Optimierungsproblem mit dem Ziel Gewinnmaximierung:

$$\begin{aligned}
 \max \quad & 16x_1 + 32x_2 \quad \text{s.t.} \quad 20x_1 + 10x_2 \leq 8000 \\
 & 4x_1 + 5x_2 \leq 2000 \\
 & 6x_1 + 15x_2 \leq 4500 \\
 & x_1, x_2 \geq 0
 \end{aligned}$$

Die Abkürzung s.t., *englisch* “subject to”, steht für “unter der Nebenbedingung” und bezeichnet die Beschränkungen des Optimierungsproblems.

Einführung von Schlupfvariablen y_1, y_2, y_3 führt auf die Beschränkungen:

$$\begin{aligned}
 20x_1 + 10x_2 + y_1 &= 8000 & y_1 &\geq 0 \\
 4x_1 + 5x_2 + y_2 &= 2000 & y_2 &\geq 0 \\
 6x_1 + 15x_2 + y_3 &= 4500 & y_3 &\geq 0
 \end{aligned}$$



In jeder Ecke sind genau 2 Ungleichungen “aktiv” (d.h. es gilt Gleichheit)
 → Löse jeweils ein lineares Gleichungssystem (LGS) mit 2 Variablen.

Optimale Lösung: $x_1 = 250$, $x_2 = 200$

Allgemeine Formulierung von linearen Optimierungsproblemen (LP)

Kanonische Form:

$$\left. \begin{array}{ll} \max & c^\top x \quad \text{über } x \in \mathbb{R}^n \\ \text{unter} & Ax \leq b \\ \text{und} & x \geq 0. \end{array} \right\} \quad (5.1)$$

Dabei sind $c \in \mathbb{R}^n$ (**Kostenvektor**), $A \in \mathbb{R}^{m \times n}$ und $b \in \mathbb{R}^m$ mit $m \geq 0$. Die Ungleichungen sind komponentenweise zu verstehen.

Lemma 5.2

Jedes LP ist äquivalent zu einem LP in kanonischer Form.

Beweis: Enthält ein gegebenes LP

- (a) eine Gleichungsbeschränkung $a^\top x = \beta$, so können wir diese als $a^\top x \leq \beta$ und $a^\top x \geq \beta$, also $-a^\top x \leq -\beta$ schreiben.
- (b) eine Beschränkung der Form $a^\top x \geq \beta$, so können wir sie mit (-1) multiplizieren.
- (c) für eine Variable x_i keine Beschränkung der Form $x_i \geq 0$ (**freie Variable**), so ersetzen wir $x_i := x_i^+ - x_i^-$ und fordern $x_i^+ \geq 0$ und $x_i^- \geq 0$.

□

Definition 5.3 (Polyeder)

Eine Teilmenge des \mathbb{R}^n , die durch endlich viele affin-lineare Gleichungs- und Ungleichungsrestriktionen beschrieben wird, heißt **Polyeder**. ◇

Jede Ungleichung $a^\top x \leq \beta$ mit $a \in \mathbb{R}^n$ und $\beta \in \mathbb{R}$ beschreibt einen Halbraum. Die zulässige Menge $P = \{x \in \mathbb{R}^n : Ax \leq b, x \geq 0\}$ ist daher der Schnitt endlich vieler Halbräume.

Für die algorithmische Behandlung von LPs sind Gleichungen geeigneter.

$$\begin{array}{ll} \min & \sum_{i=1}^n c_i x_i \\ \text{unter} & \sum_{i=1}^n a_{ji} x_i = b_j \quad j = 1, \dots, m \\ & x_1, \dots, x_n \geq 0 \end{array}$$

oder kurz:

$$\begin{array}{ll} \min & c^\top x \\ \text{unter} & Ax = b \\ & x \geq 0 \end{array}$$

Definition 5.4 (LP in Normalform (Standardform))

Ein LP der Gestalt

$$\left. \begin{array}{ll} \min & c^\top x \quad \text{über } x \in \mathbb{R}^n \\ \text{unter} & Ax = b \\ \text{und} & x \geq 0 \end{array} \right\} \quad (5.2)$$

mit $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ und $c \in \mathbb{R}^n$ heißt in **Standardform** bzw. **Normalform**. \diamond

Enthält ein gegebenes LP

- (a) eine Ungleichungsbeschränkung $a^\top x \leq \beta$, so führen wir eine zusätzliche, sogenannte **Schlupfvariable** (**slack variable**) s ein und ersetzen die Ungleichung durch

$$a^\top x + s = \beta, \quad s \geq 0.$$

- (b) freie Variablen x_i , so setzen wir wie oben $x_i := x_i^+ - x_i^-$.

Mit obigen Umformungen kann man zeigen:

Lemma 5.5

Jedes LP ist äquivalent zu einem LP in Normalform.

Wir gehen meist davon aus, dass ein LP in Normalform vorliegt und dass $1 \leq m \leq n$ gilt. [Für $m = 0$ ist die Aufgabe entweder unbeschränkt (wenn ein $c_i < 0$ ist), oder die Lösung ist $x^* = 0$ (wenn $c \geq 0$ gilt). Für $m > n$ können entweder redundante Gleichungen gestrichen werden, bis $m \leq n$ wird, oder $Ax = b$ ist unlösbar.] Die zulässige Menge

$$P = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\} \quad (5.3)$$

heißt dann ein **Polyeder in Normalform**.

Definition 5.6 (Ecke eines Polyeders)

Ein Vektor $x \in P$ heißt **Ecke** eines Polyeders P (nicht notwendig in Normalform), wenn aus

$$x = \lambda y + (1 - \lambda) z$$

für $y, z \in P$ und $\lambda \in (0, 1)$ bereits $y = z$ folgt. \diamond

Der Hauptsatz der Linearen Optimierung besagt nun: Falls das LP (5.2) eine Lösung besitzt, so gibt es eine optimale Ecke des Polyeders (5.3). Es genügt also prinzipiell, die Ecken der zulässigen Menge zu analysieren. Genau das tut der sog. *Simplex*-Algorithmus (Dantzig, 1947), der in dieser Vorlesung nicht behandelt wird. Der *Simplex*-Algorithmus stellt jedoch einen Spezialfall der Aktive-Menge-Strategie dar, die im Kontext von SQP-Verfahren in Abschnitt 7.1 vorgestellt wird. Im Kapitel Numerische Verfahren wird eine andere Klasse von Methoden zur Lösung vom LP's diskutiert, welche in engem Zusammenhang mit dem für die gesamte Optimierungstheorie linearer und nichtlinearer Probleme zentralen Konzept der Dualität steht.

Satz 5.7 (Hauptsatz der linearen Optimierung - Spezialfall)

Die zulässige Menge P (5.3) des LP (5.2) sei kompakt. Dann gibt es eine optimale Ecke von P .

Beweis: Übungsaufgabe \square

§ 5.2 Dualität

Jedem Primalproblem **(P)** ordnet man ein zugehöriges Dualproblem **(P*)** zu.

$$\begin{array}{ll} \textbf{(P)} & \min c^\top x \\ & \text{unter } Ax \geq b, \quad x \geq 0 \\ \textbf{(P*)} & \max \lambda^\top b \\ & \text{unter } \lambda^\top A \leq c, \quad \lambda \geq 0 \end{array}$$

A ist $m \times n$ -Matrix, $x \in \mathbb{R}^n$, $\lambda \in \mathbb{R}^m$, zulässige Mengen $K \subset \mathbb{R}^n$, $K^* \subset \mathbb{R}^m$.

Zur anschaulichen Motivation dieser Zuordnung:

Beispiel 5.8 (Ökonomische Bedeutung des Dualproblems)

$$\textbf{(P)} \min c^\top x \quad \text{unter } Ax \geq b, \quad x \geq 0; \quad c \in \mathbb{R}^n, b \in \mathbb{R}^m$$

x_i sei die Menge eines Lebensmittel i zum Preis c_i ($i = 1, \dots, n$). Jedes Lebensmittel enthält verschiedene Nährstoffe j ($j = 1, \dots, m$).

Optimierungsproblem: Wähle die Menge der zu kaufenden Lebensmittel unter Minimierung der **Gesamtkosten** und unter **Erfüllung des minimalen Nährstoffbedarfs** aus, der durch die Beschränkungen **$Ax \geq b$** beschrieben wird. Für jeden Nährstoff $j \in \{1, \dots, m\}$ ist eine Minimalbeschränkung der Form $\sum_{i=1}^n a_{ji}x_i \geq b_j$ zu erfüllen. a_{ji} gibt dabei den Gehalt von Lebensmittel i an Nährstoff j an.

Duales Problem: Ein Pharmaunternehmen bietet einzelne Nährstoffe in Pillenform an und möchte den Kunden überzeugen, anstelle der Lebensmittel Pillen zu kaufen, die den Nährstoffbedarf decken.

Ziel: Bestimme Preise **λ_j** für die einzelnen Nährstoffpillen $j = 1, \dots, m$ unter Gewinnmaximierung und der Beschränkung dennoch preisgünstiger für die Ernährung des Kunden zu sein als Lebensmittel.

Mathematische Formulierung:

$$\textbf{(P*)} \max \lambda^\top b \quad \text{unter } \lambda^\top a^i \leq c_i (i = 1, \dots, n), \lambda \geq 0$$

wobei a^i den i -ten Spaltenvektor der Matrix A bezeichnet.

Bedeutung von $\lambda^\top a^i \leq c_i$: Wenn man den Nährstoffgehalt von Lebensmittel i in Form von Pillen deckt, dann darf die Summe der Kosten nicht größer sein als für Lebensmittel i .

Die obige Formulierung **(P)** und **(P*)** nennt man **symmetrische Form** der Dualität. Wir betrachten nun ein LP in Normalform, also

$$\left. \begin{array}{l} \min \quad c^\top x \quad \text{über } x \in \mathbb{R}^n \\ \text{sodass} \quad Ax = b \\ \text{und} \quad x \geq 0 \end{array} \right\} \quad (5.4)$$

mit $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ und $c \in \mathbb{R}^n$.

Zu (5.4) duales LP:

$$\left. \begin{array}{ll} \max & b^\top \lambda \quad \text{über } \lambda \in \mathbb{R}^m \\ \text{sodass} & A^\top \lambda \leq c \end{array} \right\} \quad (5.5)$$

Führen wir in (5.5) die Schlupfvariablen $s \in \mathbb{R}^n$ ein, so erhalten wir die äquivalente Darstellung:

$$\left. \begin{array}{ll} \max & b^\top \lambda \quad \text{über } \lambda \in \mathbb{R}^m, s \in \mathbb{R}^n \\ \text{sodass} & A^\top \lambda + s = c \\ \text{und} & s \geq 0. \end{array} \right\} \quad (5.6)$$

Achtung: Das LP (5.6) liegt nicht in Normalform vor, da die Bedingung $\lambda \geq 0$ fehlt.

Definition 5.9 (Duales LP)

Das LP (5.5) bzw. (5.6) heißt das zu (5.4) gehörige **duale LP**. In diesem Zusammenhang heißt (5.4) das **primale LP**. Man spricht auch von **primal-dualen Paaren**.
◇

Ziel: Verständnis des Zusammenhangs von (5.4) und (5.5) bzw. (5.6)

Satz 5.10 (Schwache Dualität)

Es sei $x \in \mathbb{R}^n$ zulässig für das primale LP (5.4), und es sei $(\lambda, s) \in \mathbb{R}^m \times \mathbb{R}^n$ zulässig für das duale LP (5.6). Dann gilt für die Funktionswerte

$$b^\top \lambda \leq c^\top x.$$

Beweis: Aus der Zulässigkeit ergibt sich

$$b^\top \lambda = (Ax)^\top \lambda = x^\top (A^\top \lambda) = x^\top (c - s) \leq c^\top x,$$

denn wegen $x \geq 0$ und $s \geq 0$ gilt $x^\top s \geq 0$. □

Bemerkung 5.11 (Motivation des dualen LPs)

Jeder zulässige Punkt $A^\top \lambda \leq c$ des dualen LPs (5.5) liefert mit $b^\top \lambda$ eine untere Schranke für den optimalen Zielfunktionswert des primalen LPs. Wegen $Ax = b$ im primalen LP gilt auch $\lambda^\top Ax = \lambda^\top b$ für alle $\lambda \in \mathbb{R}^m$ (Linearkombination der Gleichungsnebenbedingungen). Im dualen LP suchen wir also eine Linearkombination der Gleichungsnebenbedingungen $Ax = b$ (repräsentiert durch λ), die den Wert der primalen Zielfunktion am stärksten einschränkt. ◇

Korollar 5.12

Es sei $x \in \mathbb{R}^n$ zulässig für das primale LP (5.4), und es sei $(\lambda, s) \in \mathbb{R}^m \times \mathbb{R}^n$ zulässig für das duale LP (5.6). Weiter gelte

$$c^\top x = b^\top \lambda.$$

Dann ist x eine Lösung des primalen LP, und (λ, s) ist eine Lösung des dualen LP.

Beweis: Seien \hat{x} und $(\hat{\lambda}, \hat{s})$ beliebige zulässige Punkte für das primale bzw. das duale LP. Aus der schwachen Dualität (Satz 5.10) folgt

$$b^\top \hat{\lambda} \leq c^\top x = b^\top \lambda \leq c^\top \hat{x},$$

d.h., x ist eine Lösung von (5.4), und (λ, s) ist Lösung von (5.6). □

Satz 5.13 (Dualitätssatz)

Besitzt eines der Probleme (5.4) oder (5.5) eine endliche Lösung, so besitzt auch das jeweils andere Problem eine endliche Lösung und es gilt für diese Lösungen x und λ :

$$\lambda^\top b = c^\top x.$$

Beweis: Es besitze o.B.d.A. das Primalproblem eine endliche Lösung x^* mit $z_0 = cx^*$. Definiere den abgeschlossenen, konvexen Kegel

$$C := \{(r, w) = t(cx - cx^*, b - Ax)^\top : x \geq 0, t > 0\} \subset \mathbb{R}^{1+m}$$

Offenbar gilt $0 \in C$ (setze $x = x^* \geq 0$).

Behauptung: $(-1, 0, \dots, 0) \notin C$.

Annahme: $(-1, 0, \dots, 0) \in C \Rightarrow \exists t_0 > 0, x_0 \geq 0$ mit $(-1, 0, \dots, 0) = t_0(cx_0 - cx^*, b - Ax_0) \Rightarrow t_0(b - Ax_0) = 0 \Rightarrow Ax_0 = b$ und $t_0(cx_0 - cx^*) = -1$. \nmid Das ist ein Widerspruch zur Annahme, denn $cx_0 \geq cx^*$ (da x^* optimal).

Es gilt $t(cx - z_0) = c(tx) - tz_0 = cy - tz_0$ mit $y := tx \geq 0$ (da $x \geq 0$).

$$t(b - Ax) = tb - A(tx) = tb - Ay \Rightarrow C = \{(cy - tz_0, tb - Ay) : t > 0, y \geq 0\}$$

Wegen $(-1, 0, \dots, 0) \notin C$ gibt es nach dem Trennungssatz (Korollar 2.30) einen Zeilenvektor $(\lambda_0, \lambda) \in \mathbb{R} \times \mathbb{R}^m$ mit $(\lambda_0, \lambda)(-1, 0, \dots, 0)^\top = -\lambda_0 < 0 \leq (\lambda_0, \lambda)(r, w)^\top \forall (r, w) \in C$, also gilt $\lambda_0 > 0$!

O.E. kann $\lambda_0 = 1$ gesetzt werden (Division durch $\lambda_0 > 0$). Nach Def. von C folgt $0 \leq (1, \lambda)(cy - tz_0, tb - Ay)^\top = cy - tz_0 + \lambda(tb - Ay) \Rightarrow 0 \leq (c - \lambda A)y - tz_0 + \lambda tb \forall t > 0, y \geq 0$. Für $t \searrow 0$ folgt $0 \leq (c - \lambda A)y \forall y \geq 0 \Rightarrow 0 \leq c - \lambda A \Rightarrow \lambda A \leq c$, also ist λ zulässig für das duale Problem P^* .

Setze $t = 1, y = 0$. Dann gilt $0 \leq -z_0 + \lambda b \Rightarrow z_0 = cx^* \leq \lambda b$. Aus dem schwachen Dualitätssatz 5.10 und 5.12 folgt die Gleichheit $\lambda b = cx^*$ und die Optimalität von λ für P^* . \square

Bemerkung 5.14

Ein alternativer Beweis mit Hilfe des Lemmas von Farkas ist möglich, er findet sich in einigen Lehrbüchern. Da das Farkas Lemma aber ebenfalls den Trennungssatz benötigt, bleibt dieser zentrales Hilfsmittel bei der Herleitung des Dualitätssatzes.

Schattenpreisformel:

Sei $z(x) = c^\top x$ die Zielfunktion der optimalen Lösung von **(P)**. Dann gibt es (wg. Dualitätssatz) ein λ als Lösung des Dualproblems mit:

$$z(x) = c^\top x = \lambda^\top b. \tag{5.7}$$

$$\text{Man kann zeigen: } z'(b) = \frac{dz}{db} = \lambda$$

Dabei ist die Ableitung nach dem Vektor b komponentenweise zu verstehen.

Die Schattenpreisformel besagt, dass die Komponenten der dualen Variable λ die Sensitivität der Zielfunktion bzgl. der Beschränkungsvektoren b beschreibt. Der Begriff "Schattenpreise" ist durch die ökonomische Bedeutung motiviert (siehe Beispiel 5.8).

Satz 5.15 (Notwendige und hinreichende Optimalitätsbedingungen)

Die folgenden Aussagen sind äquivalent:

- (a) Das primale Problem (5.4) besitzt eine Lösung x^* .
- (b) Das duale Problem (5.6) besitzt eine Lösung (λ^*, s^*) .
- (c) Die **notwendigen Optimalitätsbedingungen**

$$\left. \begin{array}{l} A^\top \lambda + s = c \\ Ax = b \\ x_i s_i = 0, \quad i = 1, \dots, n \\ x \geq 0, \quad s \geq 0 \end{array} \right\} \quad (5.8)$$

besitzen eine Lösung (x^*, λ^*, s^*) .

Beweis: bei Interesse zur Übung selbst, folgt als Spezialfall aus den allgemeinen notwendigen Optimalitätsbedingungen 1. Ordnung für nichtlineare Optimierungsprobleme! \square

Die Bedingungen in Satz 5.15 c) bezeichnet man

$$A^\top \lambda + s = c, s \geq 0: \text{ Duale Zulässigkeit}$$

$$Ax = b, x \geq 0: \text{ Primale Zulässigkeit}$$

$$x^\top s = 0, x, s \geq 0: \text{ Komplementarität}$$

Die Variablen s und λ sind die Lagrange-Multiplikatoren.

Bemerkung 5.16

Der Dualitätsbegriff lässt sich von linearen auf nichtlineare und sogar nicht-differenzierbare Optimierungsprobleme verallgemeinern und führt auf das Konzept der sog. **Lagrange-Dualität**.

§ 5.3 Innere-Punkte-Verfahren für LP

Das zu lösende LP in Standardform lautet

$$\min \quad c^\top x \quad \text{unter} \quad Ax = b, \quad x \geq 0 \quad (5.9)$$

Notwendige Optimalitätsbedingungen gemäß Satz 5.15:

$$\begin{array}{l} A^\top \lambda + s = c \\ Ax = b \\ x_i s_i = 0, \quad i = 1, \dots, n \\ x \geq 0 \\ s \geq 0 \end{array} \quad (5.10)$$

Die Komplementaritätsbedingung $x_i s_i = 0 \quad \forall i$ kann man auch schreiben als $x^\top s = 0$ (wegen $x_i \geq 0, s_i \geq 0 \quad \forall i$). Sei (x^*, λ^*, s^*) optimale Lösung von (5.9). Dann gilt:

$$\begin{aligned} c^\top x^* &= (A^\top \lambda^* + s^*) x^* = (Ax^*)^\top \lambda^* \\ &= b^\top \lambda^* \quad (\text{siehe auch Dualitätssatz 5.13}) \end{aligned}$$

Ein lokales Minimum von (5.9) ist ein globales Minimum ((5.9) ist konvex !):

Sei \bar{x} irgendein zulässiger Punkt von (5.9), d.h. $A\bar{x} = b, \bar{x} \geq 0$, dann folgt

$$c^\top \bar{x} = (A\lambda^* + s^*)^\top \bar{x} = b^\top \lambda^* + \bar{x}^\top s^* \geq b^\top \lambda^* = c^\top x^*.$$

Daher ist ein zulässiger Punkt \bar{x} von (5.9) optimal $\Leftrightarrow \bar{x}^\top s^* = 0$.

Duales Problem:

$$\begin{aligned} \max \quad & b^\top \lambda \quad \text{unter} \quad A^\top \lambda \leq c \\ & \text{bzw.} \\ \max \quad & b^\top \lambda \quad \text{unter} \quad A^\top \lambda + s = c, \quad s \geq 0 \end{aligned} \tag{5.11}$$

Umformulierung von (5.11):

$$\min \quad -b^\top \lambda \quad \text{unter} \quad c - A^\top \lambda \geq 0 \tag{5.12}$$

Sei x die duale Variable für (5.12). Notwendige Optimalitätsbedingungen (siehe KKT-Theorie im späteren Kapitel und Satz 5.15) für das Problem (5.12):

$$\begin{aligned} Ax &= b, \quad x \geq 0 \\ A^\top \lambda &\leq c, \quad x_i(c - A^\top \lambda)_i = 0, \quad i = 1, \dots, n \end{aligned} \tag{5.13}$$

Setze $s := c - A^\top \lambda$, dann sind (5.13) und (5.10) identisch!

D.h. die Lagrange-Multiplikatoren des Primalproblems sind die Optimierungsvariablen des Dualproblems und umgekehrt. Das Dualproblem des Dualproblems ist das Primalproblem.

Dualitätslücke: Sei x primal zulässig, d.h. $Ax = b, x \geq 0$ und (λ, s) dual zulässig, d.h. $A^\top \lambda + s = c, s \geq 0$. Dann: $c^\top x - b^\top \lambda = (c - A^\top \lambda)^\top x = \underbrace{s^\top x}_{\text{Duale Lücke}} \geq 0$.

Innere-Punkte-Verfahren (**IP-Verfahren**) oder „Interior Point Methods“ bewegen sich im Gegensatz zum Simplex-Verfahren im Inneren des zulässigen Polyeders eines LP zu einer Lösung (in einer Ecke der zulässigen Menge). Sie sind eine Alternative insbesondere für große LP und außerdem verallgemeinerbar auf nichtlineare Optimierungsaufgaben.

Wir betrachten im folgenden das primale LP in Normalform

$$\min \quad c^\top x \quad \text{sodass} \quad Ax = b \quad \text{und} \quad x \geq 0 \tag{5.14}$$

gemeinsam mit dem dualen LP

$$\max \quad b^\top \lambda \quad \text{sodass} \quad A^\top \lambda + s = c \quad \text{und} \quad s \geq 0 \tag{5.15}$$

und den Optimalitätsbedingungen

$$\left. \begin{array}{l} A^\top \lambda + s = c \\ Ax = b \\ x_i s_i = 0, \quad i = 1, \dots, n \\ x \geq 0, \quad s \geq 0 \end{array} \right\} \quad (5.16)$$

Wir wissen nach Satz 5.15, dass eine Lösung (x, λ, s) von (5.16) gleichzeitig Lösungen von (5.14) und (5.15) liefert. Wir stellen nun das Konzept sogenannter **primal-dualer Innere-Punkte-Verfahren** zur Lösung von (5.16) vor.

Wir betrachten folgendes gestörte Optimalitätssystem

$$\left. \begin{array}{l} A^\top \lambda + s = c \\ Ax = b \\ x_i s_i = \tau, \quad i = 1, \dots, n \\ x > 0, \quad s > 0 \end{array} \right\} \quad (5.17)$$

mit einem $\tau > 0$. Falls existent, so heißt die Abbildung

$$\tau \mapsto (x_\tau, \lambda_\tau, s_\tau)$$

auf die Lösung von (5.17) der **zentrale Pfad**, und (5.17) heißen **Zentraler-Pfad-Bedingungen (ZPB)**. Die Idee der Innere-Punkte-Verfahren besteht darin, den Pfad $\tau \searrow 0$ zu verfolgen.

Wir führen das zu (5.14) gehörige (primale) logarithmische **Barriere-Problem** ein:

$$\min \quad c^\top x - \tau \sum_{i=1}^n \log(x_i) \quad \text{sodass} \quad Ax = b \quad \text{und} \quad x > 0. \quad (5.18)$$

Das zum dualen Problem (5.15) gehörige Barriere-Problem lautet

$$\max \quad b^\top \lambda + \tau \sum_{i=1}^n \log(s_i) \quad \text{sodass} \quad A^\top \lambda + s = c \quad \text{und} \quad s > 0. \quad (5.19)$$

Den Zusammenhang zwischen (5.18), (5.19) und dem gestörten Optimalitätssystem (5.17) stellt der folgende Satz her, der ein Analogon zu Satz 5.15 für $\tau > 0$ ist.

Satz 5.17 (Notwendige und hinreichende Optimalitätsbedingungen)

Sei $\tau > 0$ gegeben. Dann sind die folgenden Aussagen äquivalent:

- (a) Das primale Barriere-Problem (5.18) besitzt eine Lösung x_τ .
- (b) Das duale Barriere-Problem (5.19) besitzt eine Lösung (λ_τ, s_τ) .
- (c) Die ZPB (5.17) besitzen eine Lösung $(x_\tau, \lambda_\tau, s_\tau)$.

Beweis: Hier ohne Beweis. Geht mit Konvexität der Zielfunktion via notwendige Optimalitätsbedingungen der Barriere-Probleme (s. KKT Theorie im späteren Kapitel). Siehe auch Geiger, Kanzow: Theorie und Numerik restringierter Optimierungsaufgaben, Springer 2002, Satz 4.1, S. 131/132 \square

Wir untersuchen jetzt die Lösbarkeit.

Definition 5.18 (Primal-dual zulässige Menge)

Die Menge

$$\mathcal{F} := \{(x, \lambda, s) : Ax = b, A^\top \lambda + s = c, x \geq 0, s \geq 0\}$$

heißt die **(primal-dual) zulässige Menge** und

$$\mathcal{F}^0 := \{(x, \lambda, s) : Ax = b, A^\top \lambda + s = c, x > 0, s > 0\}$$

die **(primal-dual) strikt zulässige Menge**. \diamond

Wegen Satz 5.17 ist $\mathcal{F}^0 \neq \emptyset$ notwendig dafür, dass (5.18) eine Lösung besitzt. Es ist aber auch hinreichend:

Satz 5.19

Die strikt zulässige Menge \mathcal{F}^0 sei nichtleer. Dann besitzt das primale Barriere-Problem (5.18) für jedes $\tau > 0$ eine Lösung x_τ .

Beweis: [Via Kompaktheit der Levelmenge des Barriereproblems und Anwendung von Korollar 2.10.] Seien $\tau > 0$ und ein $(\hat{x}, \hat{\lambda}, \hat{s}) \in \mathcal{F}^0$ gegeben. Es gilt also

$$A^\top \hat{\lambda} + \hat{s} = c, \quad A\hat{x} = b, \quad \hat{x} > 0, \quad \hat{s} > 0. \quad (*)$$

Wir bezeichnen mit

$$B_\tau(x) = c^\top x - \tau \sum_{i=1}^n \log(x_i)$$

die Zielfunktion des primalen Barriere-Problems (5.18). Wir werden zeigen, dass die Levelmenge

$$\mathcal{L}_\tau := \{x \in \mathbb{R}^n : Ax = b, x \geq 0, B_\tau(x) \leq B_\tau(\hat{x})\}$$

kompakt ist. Das Barriere-Problem (5.18) ist daher äquivalent zur Minimierung der stetigen Funktion B_τ über der kompakten Menge \mathcal{L}_τ , besitzt also eine globale Lösung.

$$\min B_\tau(x) \quad \text{über } x \in \mathcal{L}_\tau \quad (5.20)$$

[Die eigentlich benötigte strengere Bedingung $x > 0$ ergibt sich automatisch aus $B_\tau(x) \leq B_\tau(\hat{x})$ (da $B_\tau(x) \rightarrow \infty$ für $x \rightarrow 0$).]

Offenbar ist \mathcal{L}_τ abgeschlossen (als Schnitt von Urbildern abgeschlossener Mengen unter stetigen Funktionen). Für $x \in \mathcal{L}_\tau$ folgt aus (*)

$$\begin{aligned} B_\tau(x) + \tau \sum_{i=1}^n \log(x_i) &= c^\top x \\ &= c^\top x - \hat{\lambda}^\top (Ax - b) \\ &= c^\top x - x^\top A^\top \hat{\lambda} + b^\top \hat{\lambda} \\ &= c^\top x - x^\top (c - \hat{s}) + b^\top \hat{\lambda} \\ &= x^\top \hat{s} + b^\top \hat{\lambda}. \end{aligned}$$

Daher gilt

$$\begin{aligned} B_\tau(x) &\leq B_\tau(\hat{x}) \\ \Leftrightarrow x^\top \hat{s} + b^\top \hat{\lambda} - \tau \sum_{i=1}^n \log(x_i) &\leq B_\tau(\hat{x}) \\ \Leftrightarrow \sum_{i=1}^n [x_i \hat{s}_i - \tau \log(x_i)] &\leq B_\tau(\hat{x}) - b^\top \hat{\lambda} =: \text{const} \end{aligned}$$

Die Funktionen

$$x_i \mapsto [x_i \hat{s}_i - \tau \log(x_i)]$$

sind auf \mathbb{R}^+ nach unten beschränkt und konvergieren gegen ∞ für $x_i \rightarrow \infty$ (bilde 1. und 2. Ableitung!). Daher ist \mathcal{L}_τ auch beschränkt, also kompakt. \square

Satz 5.20 (Existenz des zentralen Pfades)

Die strikt zulässige Menge \mathcal{F}^0 sei nichtleer. Dann besitzen die ZPB (5.17) für jedes $\tau > 0$ eine Lösung $(x_\tau, \lambda_\tau, s_\tau)$. Dabei sind die (x, s) -Komponenten eindeutig bestimmt. Besitzt A vollen Rang, so ist auch λ_τ eindeutig.

Beweis: Nach Satz 5.19 besitzt das primale Barriere-Problem (5.18) für jedes $\tau > 0$ eine Lösung x_τ . Wegen Satz 5.17 existieren dann λ_τ und s_τ , sodass $(x_\tau, \lambda_\tau, s_\tau)$ eine Lösung von (5.17) ist. Das primale Barriere-Problem (5.18) besitzt eine strikt konvexe Zielfunktion (sieht man mit Satz 2.25), und die zulässige Menge ist konvex. Daher ist x_τ eindeutig bestimmt (Satz 2.26). Aufgrund der Bedingungen $x_i s_i = \tau$ für $i = 1, \dots, n$ ist damit auch s_τ eindeutig bestimmt.

Besitzt A vollen Rang ($m < n$ angenommen, da gleich viele oder mehr Gleichungsbeschränkungen (m) als Variablen (n) und Vollrangbedingung von A eine einelementige oder leere zulässige Menge definieren würden), dann ist AA^\top invertierbar, und aus $A^\top \lambda_\tau + s_\tau = c$ folgt

$$\lambda_\tau = (AA^\top)^{-1} A(c - s_\tau).$$

\square

Definition 5.21 (Strikt komplementäre Lösung)

Eine Lösung (x^*, λ^*, s^*) der Optimalitätsbedingungen (5.16) heißt **strikt komplementär**, wenn für alle $i = 1, \dots, n$ entweder $x_i^* = 0$ oder $s_i^* = 0$ gilt. \diamond

Satz 5.22 (Konvergenz für $\tau \searrow 0$)

Die strikt zulässige Menge \mathcal{F}^0 sei nichtleer, und es gelte $\tau_k \searrow 0$. Dann ist die Folge (x_k, s_k) beschränkt und besitzt daher eine konvergente Teilfolge. Jeder Häufungspunkt gehört zu einer strikt komplementären Lösung von (5.16).

Beweis: Es sei $(x_0, \lambda_0, s_0) \in \mathcal{F}^0$. Es gilt

$$\begin{aligned} (x_\tau - x_0)^\top (s_\tau - s_0) &= (x_\tau - x_0)^\top A^\top (\lambda_0 - \lambda_\tau) = (b - b)^\top (\lambda_0 - \lambda_\tau) = 0 \\ \Rightarrow x_0^\top s_\tau + x_\tau^\top s_0 &= \underbrace{x_\tau^\top s_\tau}_{=\tau n} + \underbrace{x_0^\top s_0}_{=:c>0} \\ \Rightarrow x_0^\top s_\tau + x_\tau^\top s_0 &\leq \tau n + c \leq \bar{\tau} n + c \quad \text{für alle } 0 < \tau \leq \bar{\tau} \\ \Rightarrow \|x_\tau\| \text{ und } \|s_\tau\| &\text{ sind beschränkt für alle } 0 < \tau \leq \bar{\tau}. \end{aligned}$$

Sei $\tau_k \searrow 0$ und $x_k := x_{\tau_k}$, $s_k := s_{\tau_k}$. Es existieren konvergente Teilfolgen

$$x_{k'} \rightarrow x^* \geq 0, \quad s_{k'} \rightarrow s^* \geq 0,$$

und es gilt $Ax_{k'} = b$, also auch $Ax^* = b$, sowie $x_{k'}^\top s_{k'} = \tau_{k'} n \searrow 0$, also auch $(x^*)^\top s^* = 0$.

Nach Satz 5.20 existiert $\lambda_{k'}$ mit $A^\top \lambda_{k'} = c - s_{k'}$, d.h., $c - s_{k'} \in \text{Bild}(A^\top)$ für alle k' . Da $\text{Bild}(A^\top)$ als Unterraum abgeschlossen ist, liegt auch $c - s^* \in \text{Bild}(A^\top)$, d.h., es existiert λ^* mit $A^\top \lambda^* + s^* = c$. Damit ist (λ^*, s^*) zulässig für das duale LP (5.6), und nach Satz 5.15 sind x^* und (λ^*, s^*) Lösungen von (5.4) bzw. (5.6).

Zur strikten Komplementarität:

$$\begin{aligned}
& (x_{k'} - x^*)^\top (s_{k'} - s^*) = (x_{k'} - x^*)^\top A^\top (\lambda^* - \lambda_{k'}) = (b - b)^\top (\lambda^* - \lambda_{k'}) = 0 \\
\Rightarrow & (x^*)^\top s_{k'} + x_{k'}^\top s^* = \underbrace{x_{k'}^\top s_{k'}}_{=\tau_{k'}n} + \underbrace{(x^*)^\top s^*}_{=0} = \tau_{k'}n \\
\Rightarrow & \sum_{i=1}^n \frac{x_i^*}{x_{k',i}} + \sum_{i=1}^n \frac{s_i^*}{s_{k',i}} = n \quad \text{wegen } s_{k',i} = \frac{\tau_{k'}}{x_{k',i}} \text{ und } x_{k',i} = \frac{\tau_{k'}}{s_{k',i}}.
\end{aligned}$$

Die Quotienten sind jeweils entweder $= 0$ oder konvergieren für $k' \rightarrow \infty$ gegen 1. Daraus folgt entweder $x_i^* = 0$ oder $s_i^* = 0$ für alle $i = 1, \dots, n$. (Man sieht das leicht mit Bsp. $n = 3$) \square

Wir wollen nun die Lösung der ZPB (5.17) mit Hilfe des Newton-Verfahrens beschreiben. Zu gegebenen $x \in \mathbb{R}^n$ und $s \in \mathbb{R}^n$ setzen wir

$$\begin{aligned}
X &:= \text{diag}(x_1, \dots, x_n) \\
S &:= \text{diag}(s_1, \dots, s_n) \\
e &:= (1, \dots, 1)^\top \in \mathbb{R}^n.
\end{aligned}$$

Dann ist (5.17) äquivalent zu

$$F_\tau(x, \lambda, s) := \begin{pmatrix} A^\top \lambda + s - c \\ Ax - b \\ XSe - \tau e \end{pmatrix} = 0 \quad \text{und} \quad x > 0, \quad s > 0. \quad (5.21)$$

Die Gleichung $F_0(x, \lambda, s) = 0$ ist nur “schwach” nichtlinear, aber es treten Komplikationen durch $(x, s) > 0$ auf. Die “Kunst” des numerischen Verfahrens ist es $(x, s) > 0$ zu behandeln.

Eine Möglichkeit zur numerischen Lösung von LPs ist die Verwendung des Newton-Verfahrens zur Lösung der ersten 3 Gleichungen der notwendigen Optimalitätsbedingungen (5.16) ($\tau = 0$) und eine Modifikation des Newton-Schritts, so dass $(x, s) > 0$ erfüllt ist (daher der Name **Innere-Punkte-Verfahren**). Der volle Schritt $(\Delta x, \Delta \lambda, \Delta s)$ verletzt in der Regel $(x, s) > 0$, daher

$$(x_k, \lambda_k, s_k) + \alpha(\Delta x_k, \Delta \lambda_k, \Delta s_k), \quad \alpha \in (0, 1].$$

Das macht häufig nur sehr kleine α möglich (\rightarrow geringer “Fortschritt” in Richtung Lösung). Daher löst man ein modifiziertes Problem mit $\tau > 0$ und kombiniert das Newton-Verfahren mit einem Homotopie-Verfahren $\tau \searrow 0$.

Die Jacobi-Matrix von $F_\tau : \mathbb{R}^{2n+m} \rightarrow \mathbb{R}^{2n+m}$ ist unabhängig von τ und ist gegeben durch

$$F'_\tau(x, \lambda, s) = \begin{pmatrix} 0 & A^\top & I \\ A & 0 & 0 \\ S & 0 & X \end{pmatrix}.$$

Satz 5.23

Sei $(x, \lambda, s) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$ gegeben mit $x > 0$ und $s > 0$. Besitzt $A \in \mathbb{R}^{m \times n}$ vollen Rang, dann ist die Jacobimatrix $F'_\tau(x, \lambda, s)$ regulär für jedes $\tau > 0$.

Beweis: Es sei $p = (p^1, p^2, p^3) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$ ein Vektor mit $F'_\tau(x, \lambda, s)p = 0$, also

$$A^\top p^2 + p^3 = 0 \quad (*)$$

$$Ap^1 = 0 \quad (**)$$

$$Sp^1 + Xp^3 = 0. \quad (***)$$

Wir multiplizieren $(*)$ von links mit $(p^1)^\top$ und erhalten mit $(**)$

$$0 = (p^1)^\top A^\top p^2 + (p^1)^\top p^3 = (p^1)^\top p^3.$$

Aus $(***)$ folgt $p^3 = -X^{-1}Sp^1$, also gilt

$$0 = -(p^1)^\top X^{-1}Sp^1.$$

Da $X^{-1}S$ positiv definit ist, folgt $p^1 = 0$ und auch $p^3 = 0$. Aus $(*)$ und der Voraussetzung $\text{rang}(A) = m$ folgt auch $p^2 = 0$, d.h., $F'_\tau(x, \lambda, s)$ ist injektiv und damit regulär. \square

Satz 5.23 rechtfertigt die Anwendung des Newton-Verfahrens auf (5.21). Der Newton-Schritt $F'_\tau(x^k, \lambda^k, s^k)(\Delta x, \Delta \lambda, \Delta s) = -F_\tau(x^k, \lambda^k, s^k)$ bei der Iterierten (x^k, λ^k, s^k) liefert

$$\begin{pmatrix} 0 & A^\top & I \\ A & 0 & 0 \\ S^k & 0 & X^k \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta \lambda \\ \Delta s \end{pmatrix} = - \begin{pmatrix} A^\top \lambda^k + s^k - c \\ Ax^k - b \\ X^k S^k e - \tau e \end{pmatrix} \quad (5.22)$$

Bemerkung 5.24

Da die ersten beiden Gleichungen in (5.21) linear in (x, λ, s) sind, erfüllt spätestens die zweite Iterierte (x^1, λ^1, s^1) diese Gleichungen. Der kritische Punkt ist also das Erfüllen der nichtlinearen Komplementaritätsbedingung.

Ein IP-Verfahren verfolgt gleichzeitig zwei Ziele:

- (a) Lösen der ZPB (5.21) für festes τ
- (b) und Erzwingen von $\tau \searrow 0$.

Wir ersetzen daher τ in (5.22) durch $\sigma_k \tau$ mit $\sigma_k \in [0, 1]$. Dabei heißt σ_k der **Zentrierungsparameter**. Für $\sigma_k = 1$ macht man einen reinen Newton-Schritt hin zum zentralen Pfad für das aktuelle τ , der uns aber nicht notwendig dichter an die Lösung für $\tau = 0$ bringt. Der Fall $\sigma_k < 1$ entspricht in (5.22) einem Newton-Schritt in Richtung eines verkleinerten τ . Im Extremfall $\sigma_k = 0$ machen wir einen reinen Schritt in Richtung des eigentlich zu lösenden Problems für $\tau = 0$.

In praktischen Verfahren wird τ nicht vorgegeben, sondern aus dem aktuellen Mittelwert der Werte von $x_i s_i$ bestimmt:

$$\mu_k := \frac{1}{n} (x^k)^\top s^k$$

μ_k heißt auch die **gewichtete Dualitätslücke**.

Der Algorithmus 5.25 lässt Freiheitsgrade in der Wahl der Zentrierungsparameters σ_k und der Schrittweite t_k . Man kann zeigen, dass man Verfahren konstruieren kann, die die gewichtete Dualitätslücke μ_k in jedem Schritt in Abhängigkeit von t_k und σ_k reduzieren.

Dazu setzen wir

$$\begin{aligned}(x^k(t), \lambda^k(t), s^k(t)) &:= (x^k, \lambda^k, s^k) + t(\Delta x, \Delta \lambda, \Delta s) \\ \mu_k(t) &:= x^k(t)^\top s^k(t)/n \\ (x^{k+1}, \lambda^{k+1}, s^{k+1}) &:= (x^k(t), \lambda^k(t), s^k(t)).\end{aligned}$$

Algorithmus 5.25 (Innere-Punkte-Verfahren)

- 1: Wähle (x^0, λ^0, s^0) und setze $k := 0$
- 2: **while** $\mu_k := (x^k)^\top s^k/n > \varepsilon$ **do**
- 3: Wähle $\sigma_k \in [0, 1]$ und löse

$$\begin{pmatrix} 0 & A^\top & I \\ A & 0 & 0 \\ S^k & 0 & X^k \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta \lambda \\ \Delta s \end{pmatrix} = - \begin{pmatrix} r_c^k \\ r_b^k \\ X^k S^k e - \sigma_k \mu_k e \end{pmatrix} \quad (5.23)$$

- 4: Setze

$$\begin{aligned}r_b^k &:= Ax^k - b, r_c^k := A^\top \lambda^k + s^k - c \\ x^{k+1} &:= x^k + t_k \Delta x \\ \lambda^{k+1} &:= \lambda^k + t_k \Delta \lambda \\ s^{k+1} &:= s^k + t_k \Delta s\end{aligned}$$

Dabei ist $t_k > 0$ eine geeignete Schrittweite, die $x^{k+1} > 0$ und $s^{k+1} > 0$ garantiert.

- 5: Setze $k := k + 1$
 - 6: **end while**
-

Lemma 5.26

Die Lösung $(\Delta x, \Delta \lambda, \Delta s)$ des linearen Gleichungssystems (5.23) besitzt folgende Eigenschaften:

- (a) $(\Delta x)^\top \Delta s = 0$
- (b) $\mu_{k+1} := \mu_k(t) = (1 - t(1 - \sigma_k))\mu_k$

Beweis: Teil (a) folgt sofort aus dem ersten Teil des Beweises von Satz 5.23 mit $p^1 = \Delta x$ und $p^3 = \Delta s$.

Für (b) betrachten wir die dritte Zeile in (5.23):

$$\begin{aligned}S^k \Delta x + X^k \Delta s &= -X^k S^k e + \sigma_k \mu_k e \quad \text{mit } \mu_k = (x^k)^\top s^k/n \\ \Rightarrow (s^k)^\top \Delta x + (x^k)^\top \Delta s &= -(1 - \sigma_k)(x^k)^\top s^k \quad \text{durch Summation}\end{aligned}$$

Daraus folgt

$$\begin{aligned}x^k(t)^\top s^k(t) &= (x^k)^\top s^k + t(s^k)^\top \Delta x + t(x^k)^\top \Delta s + t^2 \underbrace{(\Delta x)^\top \Delta s}_{=0} \\ &= (x^k)^\top s^k - t(1 - \sigma_k)(x^k)^\top s^k.\end{aligned}$$

Multiplikation mit $1/n$ liefert das Resultat. □

Satz 5.27

Es seien $\varepsilon \in (0, 1)$ und $\{(x^k, \lambda^k, s^k)\}$ eine durch den Algorithmus 5.25 erzeugte Folge. Weiter seien δ, ω und κ positive Konstanten und

- (a) $\mu_{k+1} \leq \left(1 - \frac{\delta}{n^\omega}\right) \mu_k, \quad k = 0, 1, 2, \dots$
 (b) $\mu_0 \leq \frac{1}{\varepsilon^\kappa}.$

Dann gilt

$$\mu_k \leq \varepsilon \quad \text{für alle } k \geq K := \left\lceil (1 + \kappa) \frac{n^\omega}{\delta} |\log(\varepsilon)| \right\rceil.$$

Beweis: Da die log-Funktion monoton wächst, folgt aus Bedingung (a)

$$\log \mu_{k+1} \leq \log \left(1 - \frac{\delta}{n^\omega}\right) + \log \mu_k \quad \text{für alle } k.$$

Vollständige Induktion und (b) zeigen

$$\begin{aligned} \log \mu_k &\leq k \log \left(1 - \frac{\delta}{n^\omega}\right) + \log \mu_0 \\ &\leq k \log \left(1 - \frac{\delta}{n^\omega}\right) + \kappa \log \frac{1}{\varepsilon} \\ &\leq k \left(-\frac{\delta}{n^\omega}\right) + \kappa \log \frac{1}{\varepsilon} && \text{wegen } \log(1 + \beta) \leq \beta \text{ für alle } \beta > -1 \\ &\leq \log \varepsilon && \text{für } k \geq (1 + \kappa) \frac{n^\omega}{\delta} \log \frac{1}{\varepsilon}. \end{aligned}$$

□

Es gibt verschiedene IP-Verfahren, die die wesentliche Bedingung (a) realisieren. Effizient sind sogenannte **Pfadverfolgungsverfahren**, bei denen die Iterierten (x^k, λ^k, s^k) in einer definierten Umgebung des zentralen Pfades gehalten werden. Die Größe der Umgebung sinkt mit der gewichteten Dualitätslücke μ_k gegen null.

KAPITEL 2

Numerische Verfahren für finite NLP

§ 6 Unbeschränkte Optimierung

Wir betrachten in diesem Kapitel das unrestringierte Optimierungsproblem (2.2) mit $\Omega = \mathbb{R}^n$ und $\mathcal{E} = \mathcal{I} = \emptyset$, also

$$\text{Minimiere } f(x) \quad \text{über } x \in \mathbb{R}^n.$$

Wir beschränken uns auf das Auffinden lokaler Optima. Im gesamten Kapitel 6 sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$ mindestens einmal stetig differenzierbar.

§ 6.1 Optimalitätsbedingungen

Satz 6.1 (Notwendige Bedingungen 1. Ordnung)

Sei x^* ein lokales Minimum und f stetig differenzierbar in einer Umgebung $U(x^*)$. Dann ist $\nabla f(x^*) = 0$.

Ist f stetig differenzierbar in einer Umgebung $U(x)$ und gilt $\nabla f(x) = 0$, so nennt man x einen **stationären Punkt** von f .

Beachte: Die Bedingung $\nabla f(x) = 0$ ist keinesfalls hinreichend dafür, dass x ein lokales Minimum von f ist. Betrachte etwa $f(x) = -x^2$ bei $x^* = 0$.

Satz 6.2 (Hinreichende Bedingungen 2. Ordnung)

Es sei f zweimal stetig differenzierbar in einer Umgebung $U(x^*)$, und es gelte

- (a) $\nabla f(x^*) = 0$
- (b) $\nabla^2 f(x^*)$ ist positiv definit (hinreichende Bedingung dafür ist, dass alle Eigenwerte der Hesse-Matrix $\nabla^2 f(x^*)$ positiv sind).

Dann ist x^* ein striktes lokales Minimum von f .

§ 6.2 Ein allgemeines Abstiegsverfahren

In den § 6.2–6.4 vernachlässigen wir zunächst die Optimalitätskriterien aus § 6.1 und orientieren uns nur an den Funktionswerten von f .

Definition 6.3 (Abstiegsrichtung)

Ein Vektor $d \in \mathbb{R}^n$ heißt **Abstiegsrichtung** von f im Punkt $x \in \mathbb{R}^n$, wenn es ein $\alpha > 0$ gibt mit

$$f(x + td) < f(x) \quad \text{für alle } t \in (0, \alpha). \quad \diamond$$

Idee: Suche entlang einer Abstiegsrichtung eine geeignete Schrittweite t , die f hinreichend verkleinert, und iteriere.

Lemma 6.4

Es seien $x \in \mathbb{R}^n$ und $d \in \mathbb{R}^n$ gegeben mit $\nabla f(x)^\top d < 0$. Dann ist d eine Abstiegsrichtung von f in x .

Beweis: Für die Richtungsableitung gilt

$$\delta f(x; d) = \lim_{t \searrow 0} \frac{f(x + td) - f(x)}{t} = \nabla f(x)^\top d < 0.$$

Also gibt es ein $\alpha > 0$ mit

$$\frac{f(x + td) - f(x)}{t} < 0 \quad \text{für alle } t \in (0, \alpha).$$

□

Anschaulich bedeutet $\nabla f(x)^\top d < 0$, dass der Winkel zwischen der Richtung d und dem negativen Gradienten $-\nabla f(x)$ kleiner als 90° ist.

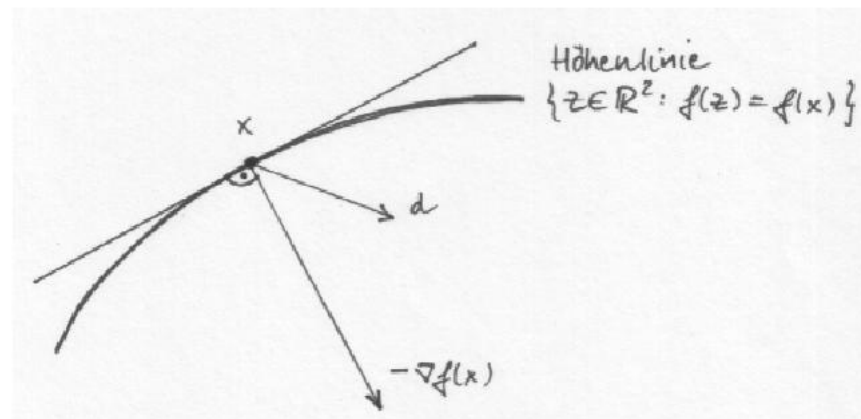


ABBILDUNG 6.1. Winkelbedingung beim Abstiegsverfahren

Beachte: Die Bedingung $\nabla f(x)^\top d < 0$ ist nicht notwendig für eine Abstiegsrichtung. Etwa in einem strikten lokalen Maximum x sind alle Richtungen $d \neq 0$ Abstiegsrichtungen, jedoch erfüllt wegen $\nabla f(x) = 0$ keine Richtung d die Bedingung $\nabla f(x)^\top d < 0$.

Bemerkung 6.5

Ist x kein stationärer Punkt von f , so ist $d = -\nabla f(x)$ immer eine Abstiegsrichtung von f in x .

Algorithmus 6.6 (Allgemeines Abstiegsverfahren)

- 1: Wähle $x^0 \in \mathbb{R}^n$ und setze $k := 0$
- 2: **while** Abbruchkriterium nicht erfüllt **do**
- 3: Bestimme eine Abstiegsrichtung d^k von f in x^k
- 4: Bestimme eine Schrittweite $t_k > 0$ mit $f(x^k + t_k d^k) < f(x^k)$
- 5: Setze $x^{k+1} := x^k + t_k d^k$ und $k := k + 1$
- 6: **end while**

Der Algorithmus lässt viele Freiheiten bei der Wahl der Abstiegsrichtungen d^k und der Schrittweiten t_k . Wir lassen zunächst das Abbruchkriterium außer Acht und wollen stets annehmen, dass der Algorithmus eine unendliche Folge $\{x^k\}$ liefert.

Definition 6.7 (Schrittweitenstrategie)

- (a) Eine Abbildung T auf $\mathbb{R}^n \times \mathbb{R}^n$, die jedem Paar (x, d) eine Teilmenge von \mathbb{R}^+ zuordnet, heißt **Schrittweitenstrategie**.
- (b) Sie heißt **wohldefiniert**, wenn $T(x, d) \neq \emptyset$ ist für alle (x, d) mit $\nabla f(x)^\top d < 0$.
- (c) Sie heißt **effizient**, wenn es für alle (x, d) mit d Abstiegsrichtung von f in x eine von x und d unabhängige Konstante $\theta > 0$ gibt mit

$$f(x + td) \leq f(x) - \theta \left(\frac{\nabla f(x)^\top d}{\|d\|} \right)^2 \quad (6.1)$$

für alle $t \in T(x, d)$.

Wir nennen eine Schrittweite $t \in T(x, d)$ selbst **effizient**, wenn T effizient ist. Ein Beispiel einer effizienten Schrittweitenstrategie werden wir in § 6.3 behandeln. Da jeweils f entlang der Richtung d^k „durchsucht“ wird, spricht man auch von Algorithmen zur **Linien suche**.

Satz 6.8 (Ein globaler Konvergenzsatz)

Es seien $\{x^k\}$, $\{d^k\}$ und $\{t_k\}$ durch Algorithmus 6.6 erzeugte Folgen mit folgenden Eigenschaften:

- (a) Die Schrittweiten $t_k > 0$ sind effizient.
- (b) Die Suchrichtungen d^k erfüllen die sogenannte **Winkelbedingung**: Es existiert eine Konstante $c > 0$ mit

$$-\nabla f(x^k)^\top d^k \geq c \|\nabla f(x^k)\| \|d^k\|. \quad (6.2)$$

Dann ist jeder Häufungspunkt x^* von $\{x^k\}$ ein stationärer Punkt von f , erfüllt also $\nabla f(x^*) = 0$.

Beweis: Wegen der Effizienz von t_k gibt es eine Konstante $\theta > 0$ mit

$$f(x^{k+1}) = f(x^k + t_k d^k) \leq f(x^k) - \theta \left(\frac{\nabla f(x^k)^\top d^k}{\|d^k\|} \right)^2$$

für alle $k \in \mathbb{N}$. Aus der Winkelbedingung (6.2) folgt somit

$$f(x^{k+1}) \leq f(x^k) - \theta c^2 \|\nabla f(x^k)\|^2. \quad (6.3)$$

Es sei nun x^* ein Häufungspunkt von $\{x^k\}$. Es gibt also eine Teilfolge $\{x^{k'}\}$ mit $x^{k'} \rightarrow x^*$, und wegen der Stetigkeit von f gilt $f(x^{k'}) \rightarrow f(x^*)$. Da $\{f(x^k)\}$ aber monoton fällt, konvergiert schon die gesamte Folge $f(x^k) \rightarrow f(x^*)$ [benutze z.B. ein Einschachtelungsargument].

Somit gilt also $f(x^{k+1}) - f(x^k) \rightarrow 0$, und aus (6.3) folgt $\nabla f(x^k) \rightarrow 0$ und damit $\nabla f(x^*) = 0$, da f eine C^1 -Funktion ist. \square

Bemerkung 6.9

Für den Winkel φ_k zwischen der Suchrichtung d^k und dem negativen Gradienten $-\nabla f(x^k)$ gilt

$$\cos \varphi_k = -\frac{\nabla f(x^k)^\top d^k}{\|\nabla f(x^k)\| \|d^k\|}.$$

Die Winkelbedingung (6.2) besagt also, dass der Winkel zwischen d^k und $-\nabla f(x^k)$ gleichmäßig von 90° weg beschränkt bleibt. \diamond

§ 6.3 Eine effiziente Schrittweitenstrategie: Armijo-Regel

Es stellt sich die Frage, wie man eine wohldefinierte und effiziente Schrittweitenstrategie im Sinne von Definition 6.7 algorithmisch umsetzen kann. Die naheliegende **Minimierungsregel**

$$\text{„Bestimme } t^k := t_{\min}^k \text{ so, dass } f(x^k + t_{\min}^k d^k) = \min_{t>0} f(x^k + t d^k) \text{ gilt“} \quad (6.4)$$

ist zwar unter bestimmten Annahmen wohldefiniert und effizient, aber wegen ihres Aufwands außer in Sonderfällen für f nicht praktikabel.

Wir behandeln hier deshalb eine andere realisierbare Schrittweitenstrategie: Die **Armijo-Regel** ist eine Bedingung, die einen hinreichenden Abstieg von f sichert. Sie ist bestimmt durch die Parameter $\sigma \in (0, 1)$ und $\beta \in (0, 1)$ und die Startschrittweite $s > 0$. Zu $(x, d) \in \mathbb{R}^n \times \mathbb{R}^n$ mit $\nabla f(x)^\top d < 0$ bestimme man

$$t_A := \max\{s\beta^l : l = 0, 1, 2, \dots\},$$

sodass

$$f(x + t_A d) \leq f(x) + \sigma t_A \nabla f(x)^\top d \quad (6.5)$$

gilt.

Man testet also die Schrittweiten $s, s\beta, s\beta^2$ etc., bis zum ersten Mal (6.5) erfüllt ist, und liefert dann $T(x, d) = \{t_A\}$ zurück („**Backtracking**“).

Zur Veranschaulichung führen wir die Liniensuchfunktion

$$\varphi(t) = f(x + t d) \quad (6.6)$$

ein. Es gilt: φ ist stetig differenzierbar in $[0, \infty)$ und

$$\varphi'(t) = \nabla f(x + t d)^\top d.$$

Also lautet die Armijo-Bedingung (6.5) alternativ

$$\varphi(t_A) \leq \varphi(0) + \sigma t_A \varphi'(0). \quad (6.7)$$

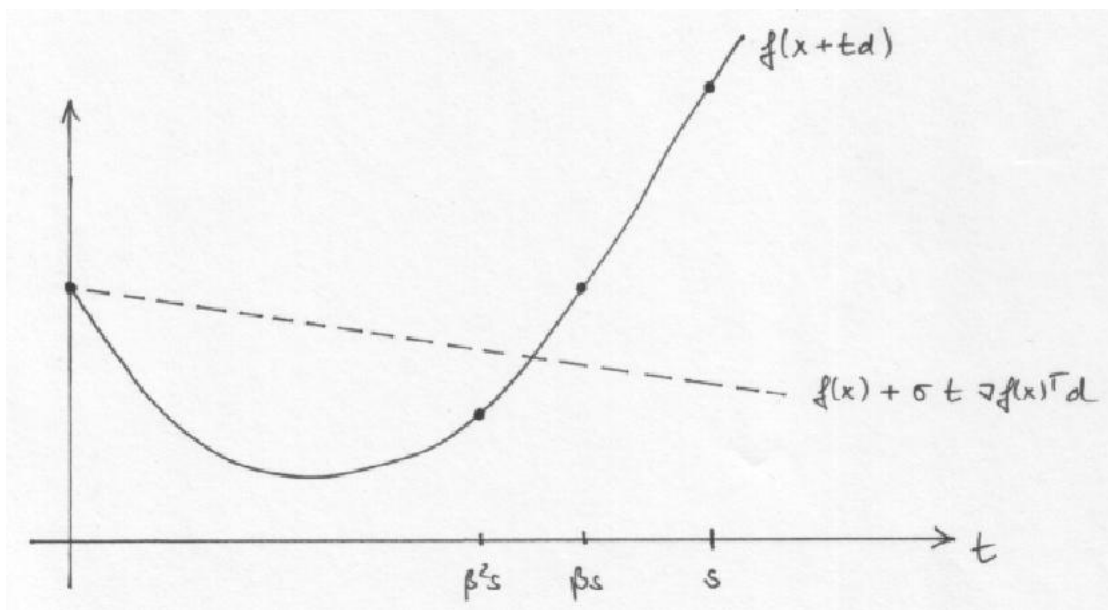


ABBILDUNG 6.2. Darstellung der Armijo-Bedingung (6.5)

Satz 6.10 (Wohldefiniertheit und Effizienz der Armijo-Regel)

Es seien $\sigma \in (0, 1)$, $\beta \in (0, 1)$ sowie $s > 0$ gegeben. Man kann unter sehr allgemeinen Bedingungen beweisen, dass die Armijo-Regel wohldefiniert und effizient ist (z.B. Lipschitz-Stetigkeit von ∇f).

§ 6.4 Wahl der Suchrichtungen: Gradientenverfahren

Bisher haben wir uns im Zusammenhang mit dem allgemeinen Abstiegsverfahren (Algorithmus 6.6) nur mit der Schrittweitenstrategie beschäftigt. In § 6.4 und § 6.5 folgen nun zwei Strategien, um die Suchrichtungen $\{d^k\}$ zu bestimmen.

Hier wählen wir zunächst

$$d = -\nabla f(x),$$

die Richtung des steilsten Abstiegs von f .

Algorithmus 6.11 (Gradientenverfahren, Verfahren des steilsten Abstiegs)

- 1: Wähle $x^0 \in \mathbb{R}^n$, $\sigma \in (0, 1)$, $\beta \in (0, 1)$, $\varepsilon > 0$ und setze $k := 0$
- 2: **while** $\|\nabla f(x^k)\| > \varepsilon$ **do**
- 3: Setze $d^k := -\nabla f(x^k)$
- 4: Bestimme eine Schrittweite t_k mit der Armijo-Regel mit der Startschrittweite $s = 1$, d.h., $t_k := \max\{\beta^l : l = 0, 1, 2, \dots\}$, sodass gilt:

$$f(x^k + t_k d^k) \leq f(x^k) + \sigma t_k \nabla f(x^k)^\top d^k$$

- 5: Setze $x^{k+1} := x^k + t_k d^k$ und $k := k + 1$
- 6: **end while**

Konvergenz bei quadratischer Zielfunktion

Um die Konvergenzgeschwindigkeit des Gradientenverfahrens zu untersuchen, wenden wir es auf eine quadratische Funktion f an:

$$f(x) = \frac{1}{2}x^\top Qx + c^\top x + \gamma$$

mit $Q \in \mathbb{R}^n \times \mathbb{R}^n$ symmetrisch positiv definit (s.p.d.), $c \in \mathbb{R}^n$ und $\gamma \in \mathbb{R}$. Die eindeutige (globale) Lösung der Aufgabe $\min_{x \in \mathbb{R}^n} f(x)$ ist

$$x^* = -Q^{-1}c, \quad (6.8)$$

denn dies ist die einzige Lösung der notwendigen Bedingungen (Satz 6.1), und die hinreichenden Bedingungen (Satz 6.2) sind dort erfüllt. Natürlich wird man das Gradientenverfahren nur anwenden, wenn die direkte Berechnung von x^* aus (6.8) zu aufwendig ist [also Q zu groß ist oder nicht explizit vorliegt].

Im Fall der quadratischen Zielfunktion lässt sich die exakte Schrittweite aus der Minierungsregel (6.4)

$$t_{min}^k = \arg \min_{t \geq 0} f(x^k + td^k)$$

im k -ten Schritt berechnen. Setzen wir

$$g^k = \nabla f(x^k) = Qx^k + c,$$

so ergibt sich

$$t_k := t_{min}^k = \frac{(g^k)^\top g^k}{(g^k)^\top Q g^k}.$$

Wir wählen also hier statt der Armijo-Regel in Algorithmus 6.11 stets die exakte Schrittweite t_{min}^k .

Es seien $\lambda_{min}(Q)$ und $\lambda_{max}(Q) > 0$ der kleinste und größte Eigenwert der s.p.d. Matrix Q und

$$\kappa = \text{cond}_2(Q) = \frac{\lambda_{max}(Q)}{\lambda_{min}(Q)}$$

die (spektrale) Konditionszahl von Q .

Satz 6.12 (Globaler Konvergenzsatz für quadratische Zielfunktionen)

Es sei Q s.p.d. Das Gradientenverfahren mit exakter Schrittweite t_{min} konvergiert für jeden Startvektor $x^0 \in \mathbb{R}^n$ gegen das eindeutige globale Minimum x^* , und es gilt

$$\begin{aligned} f(x^{k+1}) - f(x^*) &\leq \left(\frac{\kappa - 1}{\kappa + 1}\right)^2 (f(x^k) - f(x^*)) \\ \|x^k - x^*\| &\leq \sqrt{\kappa} \left(\frac{\kappa - 1}{\kappa + 1}\right)^k \|x^0 - x^*\| \end{aligned}$$

Bemerkung 6.13

Für große Konditionszahlen κ ist die Konvergenz sehr langsam (Zick-Zack-Effekt).

Zur Charakterisierung der Konvergenzgeschwindigkeit von Algorithmen führen wir ein:

Definition 6.14 (Q-Konvergenzraten)

Sei $\{x^n\} \subset \mathbb{R}^n$ eine Folge und $x^* \in \mathbb{R}^n$.

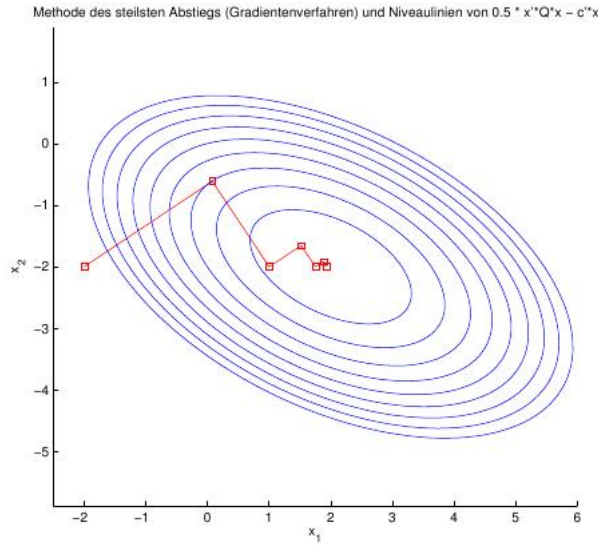


ABBILDUNG 6.3. Konvergenz des Gradientenverfahrens

- (a) $\{x^k\}$ konvergiert gegen x^* (mindestens) **q-linear**, falls ein $c \in (0, 1)$ existiert mit

$$\|x^{k+1} - x^*\| \leq c \|x^k - x^*\| \quad \text{für alle } k \in \mathbb{N} \text{ hinreichend groß}$$

- (b) $\{x^k\}$ konvergiert gegen x^* (mindestens) **q-superlinear**, falls es eine Nullfolge $\{\varepsilon_k\}$ gibt mit

$$\|x^{k+1} - x^*\| \leq \varepsilon_k \|x^k - x^*\| \quad \text{für alle } k \in \mathbb{N}.$$

- (c) Gilt $x^k \rightarrow x^*$, so konvergiert $\{x^k\}$ gegen x^* (mindestens) **q-quadratisch**, falls ein $C > 0$ existiert mit

$$\|x^{k+1} - x^*\| \leq C \|x^k - x^*\|^2 \quad \text{für alle } k \in \mathbb{N}. \quad \diamond$$

§ 6.5 Wahl der Suchrichtungen: Newton-Verfahren

[Hier orientieren wir uns jetzt wieder an den notwendigen Bedingungen!]

In diesem Abschnitt wird f als zweimal stetig differenzierbar angenommen. Wir betrachten wieder

$$\min_{x \in \mathbb{R}^n} f(x). \quad (6.9)$$

Das Newton-Verfahren lässt sich auf zwei verschiedene Weisen motivieren:

- (a) Die notwendige Optimalitätsbedingung 1. Ordnung lautet

$$\nabla f(x) = 0,$$

siehe Satz 6.1. Wende zur Lösung dieser i.A. nichtlinearen Gleichung das Newton-Verfahren an. Wir erhalten die Iterationsvorschrift

$$x^{k+1} = x^k - \nabla^2 f(x^k)^{-1} \nabla f(x^k).$$

- (b) Im aktuellen Iterationspunkt x^k ersetze man (6.9) durch die Minimierung des quadratischen Modells

$$q_k(x) = f(x^k) + \nabla f(x^k)^\top (x - x^k) + \frac{1}{2} (x - x^k)^\top \nabla^2 f(x^k) (x - x^k). \quad (6.10)$$

Ist die Hessematrix $\nabla^2 f(x^k)$ positiv definit, so ist das eindeutige Minimum durch

$$0 = \nabla q_k(x) = \nabla f(x^k) + \nabla^2 f(x^k)(x - x^k)$$

charakterisiert. Wir wählen die Lösung dieser Aufgabe als nächste Iterierte x^{k+1} und erhalten wieder

$$x^{k+1} = x^k - \nabla^2 f(x^k)^{-1} \nabla f(x^k).$$

Das Newton-Verfahren ist von der Bauart des allgemeinen Abstiegsverfahrens (Algorithmus 6.6), wobei die Suchrichtung durch die Lösung des linearen Gleichungssystems

$$\nabla^2 f(x^k) d^k = -\nabla f(x^k)$$

und die feste Schrittlänge durch $t_k = 1$ gegeben sind:

Algorithmus 6.15 (Lokales Newton-Verfahren)

- 1: Wähle $x^0 \in \mathbb{R}^n$ und $\varepsilon > 0$ und setze $k := 0$
 - 2: **while** $\|\nabla f(x^k)\| > \varepsilon$ **do**
 - 3: Löse $\nabla^2 f(x^k) d^k := -\nabla f(x^k)$
 - 4: Setze $x^{k+1} := x^k + d^k$ und $k := k + 1$
 - 5: **end while**
-

Wir können nun einen lokalen Konvergenzsatz formulieren:

Satz 6.16 (Lokaler Konvergenzsatz für Newton-Verfahren)

Es sei $x^* \in \mathbb{R}^n$ ein stationärer Punkt und $\nabla^2 f(x^*)$ regulär. Dann existiert eine Umgebung $U(x^*)$ von x^* , sodass für jedes $x^0 \in U(x^*)$ gilt:

- (a) Das lokale Newton-Verfahren ist wohldefiniert und erzeugt eine Folge $\{x^k\}$, die gegen x^* konvergiert.
- (b) Die Konvergenzrate ist q-superlinear.
- (c) Ist $\nabla^2 f$ stetig diff.bar in $U(x^*)$, so ist die Konvergenzrate q-quadratisch.

Beachte: Das vereinfachte Newton-Verfahren, bei dem statt $\nabla f(x^k)$ die feste Matrix $\nabla^2 f(x^0)$ bei der Schrittberechnung verwendet wird, konvergiert linear.

Bemerkung 6.17

Koppelt man das Newton-Verfahren an eine effiziente Schrittweitenstrategie, kann man globale Konvergenz erreichen.

§ 7 Beschränkte Nichtlineare Optimierung (NLP)

Ziel: SQP-Verfahren für allgemeine gleichungs- und ungleichungsbeschränkte nichtlineare Optimierungsprobleme.

In SQP-Verfahren wird das zu lösende NLP durch eine Folge von quadratischen Subproblemen approximiert.

§ 7.1 Quadratische Programme (QP)

Gleichungsbeschränkte QP

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) := \frac{1}{2} x^\top Q x + c^\top x + \gamma \\ \text{unter} \quad & b_j^\top x = \beta_j, \quad j = 1, \dots, p \end{aligned} \quad (7.1)$$

mit $Q \in \mathbb{R}^{n \times n}$ symmetrisch, $c \in \mathbb{R}^n$, $\gamma \in \mathbb{R}$, $b_j \in \mathbb{R}^n$, $\beta_j \in \mathbb{R}$.

Sei x^* Lösung (lokales Minimum) von (8.7), dann existieren nach dem Satz von Karush-Kuhn-Tucker Lagrange-Multiplikatoren λ_j^* , $j = 1, \dots, p$, so dass (x^*, λ^*) den KKT-Bedingungen

$$\begin{aligned} Qx + c + \sum_{j=1}^p b_j \lambda_j &= 0 \\ b_j^\top x &= \beta_j, \quad j = 1, \dots, p \end{aligned}$$

genügt.

Mit $B := (b_j^\top)_{j=1}^p = \begin{pmatrix} b_1^\top \\ \vdots \\ b_p^\top \end{pmatrix} \in \mathbb{R}^{p \times n}$ und $\beta := (\beta_1, \dots, \beta_p)^\top$ erhält man das KKT-System

$$\begin{aligned} Qx + B^\top \lambda &= -c \\ Bx &= \beta. \end{aligned}$$

Damit ergibt sich folgender Satz:

Satz 7.1

Ein Paar $(x^*, \lambda^*) \in \mathbb{R}^n \times \mathbb{R}^p$ ist genau dann ein KKT-Punkt des Optimierungsproblems (8.7), wenn (x^*, λ^*) Lösung des linearen Gleichungssystems

$$\begin{pmatrix} Q & B^\top \\ B & 0 \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} -c \\ \beta \end{pmatrix}$$

ist.

Man erhält also die Lösung eines gleichungsbeschränkten QP's mit pos. def. Matrix Q direkt als Lösung eines Linearen Gleichungssystems, da strikt konvexe NLP die KKT-Bedingungen bereits hinreichend für ein Minimum sind und ein KKT-Punkt somit einem globalen Minimum entspricht.

bei den anderen
 $\mu = 0$
nur
aktiv

$$KKT \quad L(x, \lambda, \mu) = Df(x) + g'(x)\lambda + h'(x)\mu$$

$$g(x) \leq 0, \quad \checkmark$$

$$h(x) = 0, \quad \checkmark$$

74

2. NUMERISCHE VERFAHREN FÜR FINITE NLP

$$\begin{pmatrix} Q & A & B \\ A^T & 0 & 0 \\ B^T & 0 & 0 \end{pmatrix}$$

$$\mu \geq 0$$

Ungleichungsbeschränkte QP:

Strategie der aktiven Menge (Active Set Strategy, ASS)

$$g_i(x)\mu_i = 0$$

$$\rightarrow \mu_i = 0 \text{ inaktiv}$$

$$g_i(x) = 0 \text{ aktiv}$$

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) := \frac{1}{2}x^\top Qx + c^\top x + \gamma \\ \text{unter} \quad & b_j^\top x = \beta_j, \quad j = 1, \dots, p \\ & a_i^\top x \leq \alpha_i, \quad i = 1, \dots, m \end{aligned} \quad (7.2)$$

mit $Q \in \mathbb{R}^{n \times n}$ symmetrisch, $c \in \mathbb{R}^n$, $\gamma \in \mathbb{R}$, $a_i, b_j \in \mathbb{R}^n$, $\alpha_i, \beta_j \in \mathbb{R}$.

Idee: Löse iterativ Folge von QP's unter Berücksichtigung der im aktuellen Iterationspunkt aktiven Ungleichungen.

Definiere Indexmenge der in der Lösung x^* aktiven Ungleichungen von (7.2):

$$I(x^*) := \{i \mid a_i^\top x^* - \alpha_i = 0\}$$

und approximiere diese unbekannte "aktive Menge" durch die Indexmenge \mathcal{A}_k der in der Iterierten x^k aktive Menge

$$\mathcal{A}_k := \{i \mid a_i^\top x^k - \alpha_i = 0\}.$$

Bezeichne für den folgenden Algorithmus

$$A_k := \begin{pmatrix} \vdots \\ a_i^\top \\ \vdots \end{pmatrix}_{i \in \mathcal{A}_k} \in \mathbb{R}^{\#\mathcal{A}_k \times n}$$

$$B := (b_j^\top)_{j=1}^p \in \mathbb{R}^{p \times n}.$$

Gesucht ist die aktive Menge im Lösungspunkt des QP. Ist diese bekannt, sind x, λ, μ direkt berechenbar aus einem gleichungsbeschränkten QP wie im vorherigen Abschnitt.

Algorithmus 7.2 (ASS für QP) Voraussetzung: Nebenbedingungen linear unabhängig
H pos. definit auf kern(A)

- 1: Bestimme einen für (7.2) zulässigen Punkt $x^0 \in \mathbb{R}^n$, $x^0 \in S = \{x \mid b_j^\top x - \beta_j = 0, a_i^\top x - \alpha_i \leq 0\}$, $\lambda^0 \in \mathbb{R}^p$, $\mu^0 \in \mathbb{R}^m$ beliebig. Setze $\mathcal{A}_0 := \{i \mid a_i^\top x^0 - \alpha_i = 0\}$ und $k = 0$.
- 2: Ist (x^k, λ^k, μ^k) ein KKT-Punkt von (7.2): STOP
- 3: Setze $\mu_i^{k+1} := 0$ für $i \notin \mathcal{A}_k$ und bestimme $(\Delta x^k, \lambda^{k+1}, \mu_{\mathcal{A}_k}^{k+1})$ als Lösung der Gleichungssystems

$$\begin{pmatrix} Q & A_k^\top & B^\top \\ A_k & 0 & 0 \\ B & 0 & 0 \end{pmatrix} \begin{pmatrix} \Delta x^k \\ \mu_{\mathcal{A}_k} \\ \lambda \end{pmatrix} = \begin{pmatrix} -\nabla f(x^k) \\ 0 \\ 0 \end{pmatrix} \quad (7.3)$$

wobei $\nabla f(x^k) = Qx^k + c$.

- 4: Unterscheide:

- 4.1: Ist $\Delta x^k = 0$ und $\mu_i^{k+1} \geq 0$ für alle $i \in \mathcal{A}_k$: STOP, „Lösung gefunden“.
- 4.2: Ist $\Delta x^k = 0$ und $\min\{\mu_i^{k+1}, i \in \mathcal{A}_k\} < 0$, so bestimme einen Index q mit $\mu_q^{k+1} = \min\{\mu_i^{k+1}, i \in \mathcal{A}_k\}$, setze

$$\begin{aligned} x^{k+1} &:= x^k \\ \mathcal{A}_{k+1} &:= \mathcal{A}_k \setminus \{q\} \end{aligned}$$

und gehe zu 5.

- 4.3: Ist $\Delta x^k \neq 0$ und $x^k + \Delta x^k$ zulässig für (7.2), setze

$$\begin{aligned} x^{k+1} &:= x^k + \Delta x^k \quad \text{„gehe den Schritt“} \\ \mathcal{A}_{k+1} &:= \mathcal{A}_k \quad \text{„behalte aktive Menge“} \end{aligned}$$

und gehe zu 5.

- 4.4: Ist $\Delta x^k \neq 0$, aber $x^k + \Delta x^k$ unzulässig für (7.2), so bestimme Index r mit

$$\frac{\alpha_r - a_r^\top x^k}{a_r^\top \Delta x^k} = \min \left\{ \frac{\alpha_i - a_i^\top x^k}{a_i^\top \Delta x^k} \mid i \notin \mathcal{A}_k \text{ mit } a_i^\top \Delta x^k > 0 \right\},$$

setze

$$t_k := \frac{\alpha_r - a_r^\top x^k}{a_r^\top \Delta x^k} \quad \text{„Schrittweite“}$$

$$x^{k+1} := x^k + t_k \Delta x^k \quad \text{„reduzierter Schritt“}$$

$$\mathcal{A}_{k+1} := \mathcal{A}_k \cup \{r\} \quad \text{„Aktivierung“}$$

und gehe zu 5.

- 5: $k = k + 1$ und gehe zu 2.

$$\nabla f(x) + \mu \nabla g(x) = 0$$

$$\nabla f(x) = -\mu \nabla g(x)$$

an dem Vorzeichen von μ kann man sehen ob der $\text{grad} f$ und $\text{grad} g$ in die selbe oder entgegengesetzte richtung zeigen --> ersten Folien

Erläuterungen:

- i) Die Iterierten x^k sind immer zulässig, denn x^0 ist zulässig gewählt und der Rest ($k > 0$) bleibt zulässig, da die Gleichungs und Ungleichungsbedingungen **linear** sind.
- ii) Der Vektor $(x^{k+1}, \lambda^{k+1}, \mu_{\mathcal{A}_k}^{k+1})$ aus 3. ist KKT-Punkt des gleichungsbeschränkten Problems

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & \frac{1}{2} x^\top Q x + c^\top x \\ \text{unter} \quad & b_j^\top x = \beta_j, \quad j = 1, \dots, p \\ & a_i^\top x = \alpha_i, \quad \forall i \in \mathcal{A}_k \end{aligned} \quad (7.4)$$

Das sieht man folgendermaßen, wenn man sich $a_i^\top x^k = \alpha_i, i \in \mathcal{A}_k$ zusammengefasst denkt mit $b_j^\top x^k = \beta_j, j = 1, \dots, p$ zu $B^\top x^k = \beta$.

$$\begin{aligned} & \begin{pmatrix} Q & B^\top \\ B & 0 \end{pmatrix} \begin{pmatrix} x^{k+1} \\ \mu \end{pmatrix} = \begin{pmatrix} -c \\ \beta \end{pmatrix} \\ \iff & \begin{pmatrix} Q & B^\top \\ B & 0 \end{pmatrix} \begin{pmatrix} x^k + \Delta x^k \\ \mu \end{pmatrix} = \begin{pmatrix} -c \\ \beta \end{pmatrix} \\ \iff & \begin{pmatrix} Q & B^\top \\ B & 0 \end{pmatrix} \begin{pmatrix} \Delta x^k \\ \mu \end{pmatrix} = \begin{pmatrix} -c \\ \beta \end{pmatrix} - \begin{pmatrix} Q & B^\top \\ B & 0 \end{pmatrix} \begin{pmatrix} x^k \\ 0 \end{pmatrix} \\ \iff & \begin{pmatrix} Q & B^\top \\ B & 0 \end{pmatrix} \begin{pmatrix} \Delta x^k \\ \mu \end{pmatrix} = \begin{pmatrix} -c - Qx^k \\ \beta - Bx^k \end{pmatrix} \\ \iff & \begin{pmatrix} Q & B^\top \\ B & 0 \end{pmatrix} \begin{pmatrix} \Delta x^k \\ \mu \end{pmatrix} = \begin{pmatrix} -\nabla f(x^k) \\ 0 \end{pmatrix} \end{aligned}$$

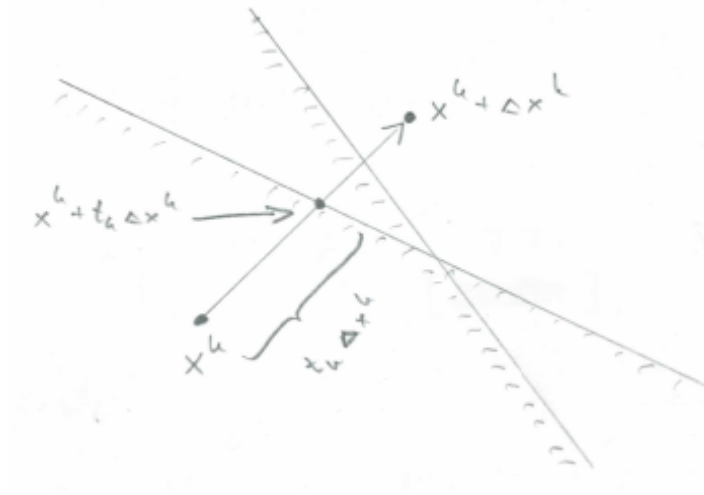
- iii) zu 4.1: Ist $\Delta x^k = 0$ und $\mu_i^{k+1} \geq 0 \forall i \in \mathcal{A}_k$ für (7.4), dann ist $(x^k, \lambda^{k+1}, \mu^{k+1})$ mit $\mu_i^{k+1} = 0 \forall i \notin \mathcal{A}_k$ auch KKT-Punkt von (7.2).
- iv) zu 4.2: Ist $\Delta x^k = 0$ und $\mu_i^{k+1} < 0$ für ein $i \in \mathcal{A}_k$, dann ist der Punkt zwar optimal für (7.4), aber nicht für (7.2). Entferne Index mit dem negativsten Multiplikator $\mu_i, i \in \mathcal{A}_k$: Inaktivierung. (Es kann auch ein anderer Index i mit negativen μ_i sein, aber genau einer !)
- v) zu 4.3: Ist $\Delta x^k \neq 0$ und $x^k + \Delta x^k$ zulässig, dann gehe Schritt Δx^k (Lösung von (7.4)) und behalte \mathcal{A}_k bei.
- vi) zu 4.4: Ist $\Delta x^k \neq 0$, aber $x^k + \Delta x^k$ unzulässig für (7.2), dann verletzt der Schritt eine bisher inaktive Ungleichungsrestriktion. Gehe also einen reduzierten Schritt $x^{k+1} := x^k + t_k \Delta x^k$ bis zum Rand der zulässigen Menge, so dass alle Ungleichungen $i \notin \mathcal{A}_k$ in x^{k+1} noch erfüllt sind. Das bedeutet

$$a_i^\top x^{k+1} = \underbrace{a_i^\top x^k}_{\leq \alpha_i} + t_k a_i^\top \Delta x^k \leq \alpha_i \quad \forall i \notin \mathcal{A}_k, \quad t_k > 0$$

Wenn $a_i^\top \Delta x^k \leq 0$, ist die Bedingung offensichtlich erfüllt.
 Im Fall $a_i^\top \Delta x^k > 0$ wähle t_k wie in 4.4, dann gilt

$$t_k \leq \frac{\alpha_i - a_i^\top x^k}{a_i^\top \Delta x^k} \quad \forall i \notin \mathcal{A}_k \text{ mit } a_i^\top \Delta x^k > 0.$$

Mache also den größten zulässigen Schritt in die berechnete Richtung Δx bis die erste bisher inaktive Ungleichung aktiv wird und nehme deren Index in die aktive Indexmenge \mathcal{A}_{k+1} auf.



Zur Wohldefiniertheit von Algorithmus 7.2:

Satz 7.3

Sei Problem (7.2) gegeben mit $Q \in \mathbb{R}^{n \times n}$ symmetrisch, $c \in \mathbb{R}^n$, $\gamma \in \mathbb{R}$, $a_i, b_j \in \mathbb{R}^n$, $i = 1, \dots, m$, $j = 1, \dots, p$. Dann gilt:

- (a) Ist Q positiv definit und sind a_i , $i \in \mathcal{A}_k$ und b_j , $j = 1, \dots, p$ linear unabhängig, so ist (7.3) eindeutig lösbar.
- (b) Sind im k -ten Schritt a_i und b_j ($i \in \mathcal{A}_k$, $j = 1, \dots, p$) linear unabhängig und tritt 4.1 nicht ein, so sind auch a_i und b_j mit $i \in \mathcal{A}_{k+1}$ und $j = 1, \dots, p$ linear unabhängig.
- (c) Ist Q positiv definit, so gilt für die in 4. berechnete Richtung Δx^k (falls $\Delta x^k \neq 0$), dass

$$\nabla f(x^k)^\top \Delta x^k < 0,$$

d.h. Δx^k ist Abstiegsrichtung.

Beweis: nur für (a): siehe Lemma 7.4. (Rest: Satz 5.4, S. 202 in “C. Kanzow, C. Geiger: Theorie und Numerik restringierter Optimierungsaufgaben, Springer, Berlin 2002”) \square

Lemma 7.4

Sei A eine $p \times n$ Matrix, $p \leq n$, $\text{rang}(A) = p$ (Vollrang). Sei weiter H positiv definit auf $\{d \mid Ad = 0\} = \ker(A)$ ($\Leftrightarrow d^\top H d > 0 \quad \forall d \neq 0, Ad = 0 \Leftrightarrow d^\top H d > 0 \quad \forall d \neq 0$).

Dann ist $\begin{pmatrix} H & A^\top \\ A & 0 \end{pmatrix}$ regulär.

Beweis: Zeige, dass $\begin{pmatrix} H & A^\top \\ A & 0 \end{pmatrix} \begin{pmatrix} d \\ \lambda \end{pmatrix} = 0$ nur die triviale Lösung besitzt.

Zeile 1 liefert

$$\begin{aligned} Hd + A^\top \lambda &= 0 \quad | \cdot d^\top \\ \Rightarrow d^\top Hd + \underbrace{d^\top A^\top}_{=0 \text{ (Zeile 2)}} \lambda &= 0 \\ \Rightarrow d^\top Hd &= 0 \xrightarrow{H \text{ p.d.}} d = 0 \Rightarrow A^\top \lambda = 0 \xrightarrow{A \text{ Vollrang}} \lambda = 0. \end{aligned}$$

□

Bemerkung 7.5

i) Satz 7.3 (c) stellt sicher, dass für Q pos. def.

$$f(x^{k+1}) < f(x^k),$$

falls $\Delta x^k \neq 0$ und $t_k \neq 0$.

ii) - Es gibt nur endlich viele Zusammensetzungen für \mathcal{A}_k .

- In jedem Schritt fällt die Zielfunktion strikt.

\Rightarrow Algorithmus liefert Lösung in endlich vielen Schritten.

iii) Die Indexmengen können zyklisieren: Problem von redundanten Ungleichungen. Dann ist $t_k = 0$. Abhilfe: Störe Ungleichungen „zufällig“

$$a_i^\top x - \alpha_i + \varepsilon_i \leq 0$$

mit ε_i so klein, dass das Problem wenig gestört wird, aber größer als der numerische Rundungsfehler. Kann im ungünstigen Fall viele neue Ecken und Schritte vom Typ 4.4 erzeugen. Besser: Gute Problemformulierung, z.B. durch Elimination redundanter Beschränkungen.

Die ASS überträgt sich konzeptionell auch auf Probleme mit nichtlinearer Zielfunktion f der Form

$$\begin{aligned} \min_x \quad & f(x) \\ \text{unter} \quad & b_j^\top x = \beta_j \\ & a_i x \leq \alpha_i \end{aligned}$$

Bei nichtlinearen Gleichungs- und Ungleichungsbedingungen werden diese in SQP-Verfahren in jedem Iterationsschritt linear approximiert und es wird ein QP der obigen Form gelöst, welches man durch eine geeignete quadratische Approximation an f erhält (aus der Taylor-Reihe 2. Ordnung).

§ 7.2 Lagrange-Newton-Iteration

Betrachte das gleichungsbeschränkte Problem

$$\begin{aligned} \min_x \quad & f(x) \\ \text{unter} \quad & h(x) = 0 \end{aligned} \tag{7.5}$$

mit $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $h : \mathbb{R}^n \rightarrow \mathbb{R}^p$, f, h zweimal stetig differenzierbar. Die KKT-Bedingungen lassen sich schreiben als $\phi(x, \lambda) = 0$ mit $\phi : \mathbb{R}^n \times \mathbb{R}^p \rightarrow \mathbb{R}^n \times \mathbb{R}^p$,

$$\phi(x, \lambda) := \begin{pmatrix} \nabla_x \mathcal{L}(x, \lambda) \\ h(x) \end{pmatrix}$$

$$\nabla_x \mathcal{L}(x, \lambda) = \nabla f(x) + \sum_{j=1}^p \lambda_j \nabla h_j(x)$$

$$\mathcal{L}(x, \lambda) = f(x) + \sum_{j=1}^p \lambda_j h_j(x).$$

Es handelt sich um ein nichtlineares Gleichungssystem in $n + p$ Variablen (x, λ) . Löse Nullstellengleichung der Lagrange-Funktion mit Hilfe des Newton-Verfahrens \rightarrow Lagrange-Newton-Iteration:

Algorithmus 7.6 (Lagrange-Newton-Verfahren)

- 1: Wähle $(x^0, \lambda^0) \in \mathbb{R}^n \times \mathbb{R}^p$, setze $k := 0$
- 2: Ist $\phi(x^k, \lambda^k) = 0$: STOP
- 3: Berechne $(\Delta x^k, \Delta \lambda^k) \in \mathbb{R}^n \times \mathbb{R}^p$ als Lsg. des LGS

$$\phi'(x^k, \lambda^k) \begin{pmatrix} \Delta x^k \\ \Delta \lambda^k \end{pmatrix} = -\phi(x^k, \lambda^k)$$

- 4: Setze $(x^{k+1}, \lambda^{k+1}) := (x^k, \lambda^k) + (\Delta x^k, \Delta \lambda^k)$, $k = k + 1$ und gehe zu 2.
-

Wenn ϕ' regulär ist ergibt sich lokal superlineare bzw. lokal quadratische Konvergenz.

Satz 7.7

Sei $(x^*, \lambda^*) \in \mathbb{R}^n \times \mathbb{R}^p$ ein KKT-Punkt von (7.5), sodass die folgenden Bedingungen erfüllt sind.

- (a) Die Gradienten $\nabla h_1(x^*), \dots, \nabla h_p(x^*)$ sind linear unabhängig (LICQ).
- (b) Es gilt $d^\top \nabla_{xx}^2 \mathcal{L}(x^*, \lambda^*) d > 0$ für alle $d \neq 0$ mit $\nabla h_j(x^*)^\top d = 0$ ($j = 1, \dots, p$) (Hinreichende Bed. 2. Ordnung)

Dann ist die Jacobi-Matrix $\phi'(x^*, \lambda^*)$ regulär.

Beweis: Lemma 7.4, denn

$$\phi'(x^*, \lambda^*) = \begin{pmatrix} \nabla_{xx}^2 \mathcal{L}(x^*, \lambda^*) & \nabla_x h(x^*) \\ \nabla_x h(x^*)^\top & 0 \end{pmatrix}$$

entspricht $\begin{pmatrix} H & A^\top \\ A & 0 \end{pmatrix}$, wobei $\nabla_x h(x^*) = (\nabla_x h_1(x^*), \dots, \nabla_x h_p(x^*)) \in \mathbb{R}^{n \times p}$ □

§ 7.3 Lokales SQP-Verfahren

Betrachte wieder das Problem

$$\begin{aligned} \min_x \quad & f(x) \\ \text{unter} \quad & h(x) = 0 \end{aligned}$$

mit f, h zweimal stetig differenzierbar.

Das Lagrange-Newton-Verfahren löst

$$\phi'(x^k, \lambda^k) \begin{pmatrix} \Delta x^k \\ \Delta \lambda^k \end{pmatrix} = -\phi(x^k, \lambda^k) \quad (7.6)$$

$$\text{mit } \phi = (\nabla_x \mathcal{L}, h)^\top \text{ und } \phi' = \begin{pmatrix} \nabla_{xx}^2 \mathcal{L} & (h')^\top \\ h' & 0 \end{pmatrix}$$

und definiert $x^{k+1} := x^k + \Delta x^k$ und $\lambda^{k+1} := \lambda^k + \Delta \lambda^k$.

Das System (7.6) lässt sich auch schreiben als

$$\begin{aligned} H_k \Delta x^k + h'(x^k)^\top \Delta \lambda^k &= -\nabla_x \mathcal{L}(x^k, \lambda^k) = -\nabla f(x^k) - h'(x^k)^\top \lambda^k \\ h'(x^k) \Delta x^k &= -h(x^k), \end{aligned}$$

mit $H_k = \nabla_{xx}^2 \mathcal{L}(x^k, \lambda^k)$.

Mit $\lambda^{k+1} := \lambda^k + \Delta \lambda^k$ ist (7.6) wieder äquivalent zu

$$\begin{aligned} H_k \Delta x^k + h'(x^k)^\top \lambda^{k+1} &= -\nabla f(x^k) \\ h'(x^k) \Delta x^k &= -h(x^k). \end{aligned}$$

Das sind gerade die KKT-Bedingungen des Systems

$$\begin{aligned} \min_{\Delta x^k} \quad & \nabla f(x^k)^\top \Delta x^k + \frac{1}{2} (\Delta x^k)^\top H_k \Delta x^k \\ \text{unter} \quad & h(x^k) + h'(x^k) \Delta x^k = 0 \end{aligned}$$

Für das allgemeine NLP

$$\begin{aligned} \min_x \quad & f(x) \\ \text{unter} \quad & g(x) \leq 0 \\ & h(x) = 0 \end{aligned} \quad (7.7)$$

ergibt sich also die Idee, das Inkrement Δx^k für den Schritt $x^{k+1} = x^k + \Delta x^k$ als Lösung einer lokalen quadratischen Approximation zu berechnen:

$$\begin{aligned} \min_{\Delta x^k} \quad & \nabla f(x^k)^\top \Delta x^k + \frac{1}{2} (\Delta x^k)^\top H_k \Delta x^k \\ \text{unter} \quad & g(x^k) + g'(x^k) \Delta x^k \leq 0 \\ & h(x^k) + h'(x^k) \Delta x^k = 0. \end{aligned} \quad (7.8)$$

Hier sind die Beschränkungen linearisiert und die Zielfunktion wird quadratisch approximiert mit $H_k \approx \nabla_{xx}^2 \mathcal{L}(x^k, \mu^k, \lambda^k)$. In jedem Iterationsschritt ist ein QP zu lösen, deshalb spricht man von *Successive Quadratic Programming* oder *Sequential Quadratic Programming (SQP)*.

Algorithmus 7.8 (SQP-Verfahren)

- 1: Wähle $(x^0, \mu^0, \lambda^0) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p$, $H_0 \in \mathbb{R}^{n \times n}$ symmetrisch, $k := 0$
- 2: Ist (x^k, μ^k, λ^k) KKT-Punkt von (7.7): STOP
- 3: Berechne $\Delta x^k \in \mathbb{R}^n$ als Lösung des Sub-QP (7.8) mit Lagrange-Multiplikatoren μ^{k+1}, λ^{k+1}
- 4: Setze $x^{k+1} := x^k + \Delta x^k$, wähle $H^{k+1} \in \mathbb{R}^{n \times n}$ symmetrisch, $k = k + 1$ und gehe zu 2.

Hierbei kann H_k geeignet gewählt werden, es wird i.d.R. positive Definitheit sicher gestellt, um eine eindeutige Lösung des Sub-QP zu garantieren.

Wählt man $H_k = \nabla_{xx}^2 \mathcal{L}(x^k, \mu^k, \lambda^k) \forall k \in \mathbb{N}$, so wird das Verfahren lokal superlinear bzw. lokal quadratisch gegen die Lösung, d.h. einen KKT-Punkt von (7.7) konvergieren. Allerdings muss $\nabla_{xx}^2 \mathcal{L}$ nicht auf ganz \mathbb{R}^n pos. def. sein.

Algorithmus 7.9 (SQP-Verfahren mit $H_k = \nabla_{xx}^2 \mathcal{L}(x^k, \mu^k, \lambda^k)$)

- 1: Wähle (x^0, μ^0, λ^0) und setze $k := 0$
- 2: Ist (x^k, μ^k, λ^k) KKT-Punkt von (7.7): STOP
- 3: Berechne einen KKT-Punkt $(x^{k+1}, \mu^{k+1}, \lambda^{k+1})$ als Lösung des QP

$$\begin{aligned} \min_x \quad & \nabla f(x^k)(x - x^k) + \frac{1}{2}(x - x^k)^\top \nabla_{xx}^2 \mathcal{L}(x^k, \mu^k, \lambda^k)(x - x^k) \\ \text{unter} \quad & g(x^k) + g'(x^k)(x - x^k) \leq 0 \\ & h(x^k) + h'(x^k)(x - x^k) = 0. \end{aligned}$$

Besitzt das QP mehrere KKT-Punkte, so wähle $(x^{k+1}, \mu^{k+1}, \lambda^{k+1})$ derart, dass der Abstand $\|(x^{k+1}, \mu^{k+1}, \lambda^{k+1}) - (x^k, \mu^k, \lambda^k)\|$ minimal wird.

- 4: $k = k + 1$ und gehe zu 2.

Im Zusammenhang mit Algorithmus 7.9 lässt sich folgender Satz formulieren:

Satz 7.10

Sei $(x^*, \mu^*, \lambda^*) \in \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^m$ ein KKT-Punkt von (7.7), für den folgende Voraussetzungen gelten:

- (i) Es gilt $g_i(x^*) + \mu_i \neq 0$ für alle $i = 1, \dots, m$ (strikte Komplementarität)
- (ii) Die Gradienten $\nabla h_j(x^*)$, $j = 1, \dots, p$ und $\nabla g_i(x^*)$, $i \in I(x^*) := \{i \mid g_i(x^*) = 0\}$ sind linear unabhängig (LICQ)
- (iii) Es gilt, dass $d^\top \nabla_{xx}^2 \mathcal{L}(x^*, \mu^*, \lambda^*)d > 0$ für alle $d \neq 0$ mit $\nabla h_j(x^*)d = 0$, $j = 1, \dots, p$ und $\nabla g_i(x^*)^\top d = 0$, $i \in I(x^*)$, ist (Hinreichende Bedingung 2. Ordnung)

Dann existiert eine Umgebung \mathcal{U}_ε von (x^*, μ^*, λ^*) , so dass für jeden Startvektor $(x^0, \mu^0, \lambda^0) \in \mathcal{U}_\varepsilon(x^*, \mu^*, \lambda^*)$ und jede durch Algorithmus 7.8 erzeugte Folge $(x^k, \mu^k, \lambda^k)_{k \in \mathbb{N}}$ gilt:

- (a) Das SQP-Verfahren (Alg. 7.8) ist wohldefiniert und die Folge $(x^k, \mu^k, \lambda^k)_k$ konvergiert gegen (x^*, μ^*, λ^*) .
- (b) Die Konvergenzrate ist superlinear.
- (c) Sind $\nabla^2 f$, $\nabla^2 g_i$, $i = 1, \dots, m$ und $\nabla^2 h_j$, $j = 1, \dots, p$ lokal Lipschitz-stetig, so ist die Konvergenzrate quadratisch.

Beweis: Satz 5.31, S. 245 in “C. Kanzow, C. Geiger: Theorie und Numerik restrukturierter Optimierungsaufgaben, Springer, Berlin 2002” \square

Bemerkung 7.11

Unter Bedingung (i) sind die in x^* inaktiven Beschränkungen in einer Umgebung vernachlässigbar (stetige Abhängigkeit!). Das Verfahren ist dann lokal äquivalent zum Lagrange-Newton-Verfahren für das gleichungsbeschränkte Problem

$$\begin{aligned} \min_x \quad & f(x) \\ \text{unter} \quad & h(x) = 0 \\ & g_i(x) = 0, \quad i \in I(x^*). \end{aligned}$$

Die Voraussetzungen (ii) und (iii) garantieren wohldefinierte Lagrange-Newton-Iterationen und die Konvergenzrate des Newton-Verfahren überträgt sich.

§ 7.4 Globalisierung von SQP - Verfahren

Betrachte das allgemeine Optimierungsproblem

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & x \in S = \{x \in \mathbb{R}^n \mid g(x) \leq 0, h(x) = 0\} \end{aligned} \quad (7.9)$$

$$g : \mathbb{R}^n \rightarrow \mathbb{R}^m, h : \mathbb{R}^n \rightarrow \mathbb{R}^p.$$

Definiere eine sog. Penalty-Funktion

$$P_r(x; \alpha) = f(x) + \alpha r(x) \quad (7.10)$$

mit $r : \mathbb{R}^n \rightarrow \mathbb{R}$ stetig, $r(x) \geq 0 \quad \forall x \in \mathbb{R}^n$, $r(x) = 0 \Leftrightarrow x \in S$. r “bestraft” also ein Verlassen der zulässigen Menge.

Definition 7.12

Eine Penalty-Funktion $P_r(x; \alpha)$ heißt **exakt** in einer lokalen Minimalstelle x^* von (7.9), wenn es $0 < \bar{\alpha} < \infty$ gibt, so dass x^* lokale Minimalstelle von $P_r(x; \alpha) \quad \forall \alpha \geq \bar{\alpha}$ ist.

Bemerkung 7.13

Falls $\nabla f(x^*) \neq 0$, ist eine exakte Penalty-Funktion der Form (7.10) in x^* nicht differenzierbar.

Wähle z.B. $r_1(x) := \|g(x)_+, h(x)\|_1$ mit $g(x)_+ := \max\{0, g(x)\}$ und

$$P_1(x; \alpha) := f(x) + \alpha \|g(x)_+, h(x)\|_1 = f(x) + \alpha \sum_{i=1}^m \max\{0, g_i(x)\} + \alpha \sum_{j=1}^p |h_j(x)|$$

(exakte ℓ_1 - Penalty - Funktion)

Man kann die Konvergenz eines SQP-Verfahrens mit Hilfe einer Schrittweitenstrategie (z.B. Armijo) für $P_1(x, \alpha)$ globalisieren. $P_1(x, \alpha)$ ist nicht überall differenzierbar, aber die Richtungsableitung

$$P'_1(x, d, \alpha) = \lim_{t \searrow 0} \frac{P_1(x + td; \alpha) - P_1(x; \alpha)}{t}$$

existiert für alle $d \in \mathbb{R}^n$.

Algorithmus 7.14 (Globales SQP-Verfahren)

- 1: Wähle $(x^0, \mu^0, \lambda^0) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p$, $H_0 \in \mathbb{R}^{n \times m}$ symmetrisch, $\alpha \geq 0, \beta \in (0, 1), \sigma \in (0, 1), k := 0$
- 2: Falls (x^k, μ^k, λ^k) KKT-Punkt: STOP
- 3: Berechne Lösung $\Delta x^k \in \mathbb{R}^n$ des Sub-QP:

$$\begin{aligned} \min_{\Delta x^k} \quad & \nabla f(x^k)^\top \Delta x^k + \frac{1}{2} (\Delta x^k)^\top H_k \Delta x^k \\ \text{s.t.} \quad & g_i(x^k) + \nabla g_i(x^k)^\top \Delta x^k \leq 0 \quad i = 1, \dots, m \\ & h_j(x^k) + \nabla h_j(x^k)^\top \Delta x^k = 0 \quad j = 1, \dots, p \end{aligned}$$

$\beta = [0, 1]$

- und zugehörige Lagrange-Multiplikatoren μ^{k+1}, λ^{k+1} . Falls $\Delta x^k = 0$: STOP
- 4: Bestimme Schrittweite $t_k = \max\{\beta^l \mid l = 0, 1, \dots\}$ mit $P_1(x^k + t_k \Delta x^k; \alpha) \leq P_1(x^k; \alpha) + \sigma t_k P'_1(x^k, \Delta x^k; \alpha)$ (Armijo-Regel)
 - 5: $x^{k+1} := x^k + t_k \Delta x^k$, wähle $H_{k+1} \in \mathbb{R}^{n \times n}$ symmetrisch, $k = k + 1$ und gehe zu 2.
-

Globale Konvergenz ist beweisbar für Algorithmus (7.14). In der Praxis benutzt man in der Regel kein festes α , sondern Updates

$$\alpha_{k+1} := \max\{\alpha_k, \max\{\mu_1^{k+1}, \dots, \mu_m^{k+1}, \lambda_1^{k+1}, \dots, \lambda_p^{k+1}\} + \gamma\}, \gamma > 0 \text{ fest.}$$

§ 7.5 Innere Punkte Verfahren (IP) für NLP

Analog LP (siehe Folien NLP).

§ 7.6 Zusammenfassung der Theorie und Numerik differenzierbarer beschränkter Optimierungsprobleme

Siehe Folien NLP !

Nichtdifferenzierbare finite Optimierung

§ 8 Lagrange-Dualität und Sattelpunkte

§ 8.1 Lagrange-Dualität

Bei linearen Programmen in § 5 hatten wir ein duales LP angegeben, dessen Funktionswerte untere Schranken für das primale LP lieferten, siehe Satz 5.10.

Frage: Kann man das auf nichtlineare Optimierungsaufgaben (4.1) verallgemeinern?

Wir betrachten wieder

$$\left. \begin{array}{ll} \text{Minimiere} & f(x) \quad \text{über } x \in \Omega \\ \text{sodass} & g_i(x) \leq 0, \quad i = 1, \dots, k \\ \text{und} & h_j(x) = 0, \quad j = 1, \dots, m - k. \end{array} \right\} \quad (8.1)$$

mit $0 \leq k \leq m$, wobei nun f , g und h weder konvex noch differenzierbar sein müssen. Außerdem lassen wir als Grundmenge $\Omega \subset \mathbb{R}^n$ (echte Zeilmenge) zu, mit der typischerweise „schwierige Restriktionen“ erfasst werden, die sich nicht einfach durch Ungleichungen $g_i(x) \leq 0$ und Gleichungen $h_j(x) = 0$ beschreiben lassen, z.B. $\Omega = \mathbb{Z}^n$ (Ganzzahligkeit).

Wir definieren wieder die **Lagrange-Funktion** zu (8.1) als

$$\begin{aligned} \mathcal{L}(x, \lambda, \mu) &:= f(x) + \sum_{i=1}^k \mu_i g_i(x) + \sum_{j=1}^{m-k} \lambda_j h_j(x) \\ &= f(x) + \mu^\top g(x) + \lambda^\top h(x), \end{aligned} \quad (8.2)$$

wobei die abstrakte Restriktion $x \in \Omega$ nicht eingeht.

Definition 8.1

(a) Die Funktion

$$q(\lambda, \mu) := \inf_{x \in \Omega} \mathcal{L}(x, \lambda, \mu)$$

heißt die **(Lagrange-)duale Funktion** von (8.1).

(b) Die Optimierungsaufgabe

$$\left. \begin{array}{ll} \text{Maximiere} & q(\lambda, \mu) \quad \text{über } (\lambda, \mu) \in \mathbb{R}^{m-k} \times \mathbb{R}^k \\ \text{sodass} & \mu \geq 0 \end{array} \right\} \quad (8.3)$$

heißt die zu (8.1) **(Lagrange-)duale Optimierungsaufgabe**. In diesem Zusammenhang heißt (8.1) auch die **primale Aufgabe**. \diamond

Beachte: Das duale Problem (8.3) hat sehr einfache Restriktionen, aber evtl. eine komplizierte Zielfunktion.

Beispiel 8.2 (Lagrange-Dualität bei LP)

Wir betrachten das LP in Normalform (5.4), also

$$\begin{aligned} f(x) &= c^\top x \\ g(x) &= -x \\ h(x) &= b - Ax \end{aligned}$$

und Grundmenge $\Omega = \mathbb{R}^n$. Die Lagrangefunktion ist

$$\begin{aligned} \mathcal{L}(x, \lambda, \mu) &= c^\top x - \mu^\top x + \lambda^\top (b - Ax) \\ &= (c - \mu - A^\top \lambda)^\top x + \lambda^\top b. \end{aligned}$$

Wir berechnen

$$\begin{aligned} q(\lambda, \mu) &= \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda, \mu) \\ &= \begin{cases} -\infty, & \text{falls } c - \mu - A^\top \lambda \neq 0 \\ \lambda^\top b, & \text{falls } c - \mu - A^\top \lambda = 0. \end{cases} \end{aligned}$$

Bei der Maximierung von $q(\lambda, \mu)$ in (8.3) können wir uns deshalb auf die Paare (λ, μ) mit

$$A^\top \lambda + \mu = c \quad (*)$$

beschränken. [Die Zielfunktion $q(\lambda, \mu)$ enthält $(*)$ als implizite Beschränkung, die wir jetzt explizit mit aufnehmen.] Damit ist (8.3) äquivalent zu

$$\begin{aligned} &\text{Maximiere } b^\top \lambda && \text{über } (\lambda, \mu) \in \mathbb{R}^{m-k} \times \mathbb{R}^k \\ &\text{sodass } A^\top \lambda + \mu = c \\ &\text{und } \mu \geq 0. \end{aligned}$$

Das ist das bereits bekannte duale LP aus (5.6) [mit μ statt s]! \diamond

Das Beispiel 8.2 zeigt, dass die duale Zielfunktion $q(\lambda, \mu)$ den Wert $-\infty$ annehmen kann. Deshalb definieren wir den **eigentlichen Definitionsbereich**

$$\text{dom}(q) = \{(\lambda, \mu) \in \mathbb{R}^{m-k} \times \mathbb{R}^k : \mu \geq 0 \text{ und } q(\lambda, \mu) > -\infty\},$$

mit dem wir das duale Problem (8.3) äquivalent als

$$\text{Maximiere } q(\lambda, \mu) \text{ über } (\lambda, \mu) \in \text{dom}(q) \quad (8.4)$$

schreiben können. Dabei heißt $\text{dom}(q)$ auch die **dual zulässige Menge**.

Satz 8.3 (Eigenschaften der dualen Funktion)

Für die duale Funktion $q(\lambda, \mu)$ gilt:

- (a) Die Menge $\text{dom}(q)$ ist konvex.
- (b) Die Funktion $q : \text{dom}(q) \rightarrow \mathbb{R}$ ist konkav.

Beachte: Dies gilt ohne Konvexitätsannahmen für das primale Problem!

Beweis: Es sei $x \in \Omega$ beliebig und $(\lambda_1, \mu_1), (\lambda_2, \mu_2) \in \text{dom}(q)$ sowie $\alpha \in [0, 1]$.

$$\begin{aligned}
 & \mathcal{L}(x, \alpha \lambda_1 + (1 - \alpha) \lambda_2, \alpha \mu_1 + (1 - \alpha) \mu_2) \\
 &= f(x) + [\alpha \mu_1 + (1 - \alpha) \mu_2]^\top g(x) + [\alpha \lambda_1 + (1 - \alpha) \lambda_2]^\top h(x) \\
 &= \alpha [f(x) + \mu_1^\top g(x) + \lambda_1^\top h(x)] + (1 - \alpha) [f(x) + \mu_2^\top g(x) + \lambda_2^\top h(x)] \\
 &= \alpha \mathcal{L}(x, \lambda_1, \mu_1) + (1 - \alpha) \mathcal{L}(x, \lambda_2, \mu_2).
 \end{aligned}$$

Daher gilt

$$\begin{aligned}
 & q(\alpha \lambda_1 + (1 - \alpha) \lambda_2, \alpha \mu_1 + (1 - \alpha) \mu_2) \\
 &= \inf_{x \in \Omega} \mathcal{L}(x, \alpha \lambda_1 + (1 - \alpha) \lambda_2, \alpha \mu_1 + (1 - \alpha) \mu_2) \\
 &\geq \alpha \inf_{x \in \Omega} \mathcal{L}(x, \lambda_1, \mu_1) + (1 - \alpha) \inf_{x \in \Omega} \mathcal{L}(x, \lambda_2, \mu_2) \\
 &= \alpha q(\lambda_1, \mu_1) + (1 - \alpha) q(\lambda_2, \mu_2).
 \end{aligned}$$

Die rechte Seite ist $> -\infty$, also auch die linke, d.h., $\alpha(\lambda_1, \mu_1) + (1 - \alpha)(\lambda_2, \mu_2) \in \text{dom}(q)$, damit ist $\text{dom}(q)$ konvex. Außerdem bedeutet die Ungleichung, dass q auf $\text{dom}(q)$ konkav ist. \square

Korollar 8.4

Das zum dualen Problem (8.3) bzw. (8.4) äquivalente Problem

$$\text{Minimiere } -q(\lambda, \mu) \text{ über } (\lambda, \mu) \in \text{dom}(q) \quad (8.5)$$

ist immer ein konvexes Optimierungsproblem, vgl. (2.10). Insbesondere sind lokale Optima bereits globale Optima (Satz 2.26). Die zulässige Menge $\text{dom}(q)$ kann jedoch leer sein.

Satz 8.5 (Schwache Dualität)

Es sei $x \in \mathbb{R}^n$ zulässig für das primale Problem (8.1), und es sei $(\lambda, \mu) \in \mathbb{R}^{m-k} \times \mathbb{R}^k$ zulässig für das duale Problem (8.3). Dann gilt für die Funktionswerte

$$q(\lambda, \mu) \leq f(x).$$

Beachte: Das duale Problem liefert untere Schranken an den Zielfunktionswert der primalen Aufgabe (vgl. Satz 5.10 und Bemerkung 5.11)! Mit der dualen Aufgabe (8.4) wird also die beste (größte) untere Schranke bestimmt.

Beweis: Es gilt

$$\begin{aligned}
 q(\lambda, \mu) &= \inf_{x \in \Omega} \mathcal{L}(x, \lambda, \mu) \\
 &\leq \mathcal{L}(x, \lambda, \mu) \\
 &= f(x) + \sum_{i=1}^k \underbrace{\mu_i}_{\geq 0} \underbrace{g_i(x)}_{\leq 0} + \sum_{j=1}^{m-k} \lambda_j \underbrace{h_j(x)}_{=0} \\
 &\leq f(x).
 \end{aligned}$$

\square

Wir setzen

$$\begin{aligned}
 \inf(P) &:= \inf\{f(x) : x \in \Omega, g(x) \leq 0, h(x) = 0\} \\
 \sup(D) &:= \sup\{q(\lambda, \mu) : (\lambda, \mu) \in \text{dom}(q)\}.
 \end{aligned}$$

Satz 8.5 bedeutet:

$$\sup(D) \leq \inf(P).$$

Dies gilt auch für die möglichen unbeschränkten/unzulässigen Fälle $-\infty \leq -\infty$, $-\infty \leq \infty$ und $\infty \leq \infty$. Anders als bei LP kann im allgemeinen Fall eine **Dualitätslücke**, d.h. $\sup(D) < \inf(P)$, auftreten, siehe nachfolgendes Beispiel:

Beispiel 8.6

Betrachte die Aufgabe

$$\begin{aligned} \text{Minimiere } f(x) &:= \begin{cases} x^2 - 2x, & \text{falls } x \geq 0 \\ x, & \text{falls } x < 0 \end{cases} \quad \text{über } x \in \Omega = \mathbb{R} \\ \text{sodass } g(x) &:= -x \leq 0. \end{aligned}$$

Es gilt

$$\mathcal{L}(x, \mu) = \begin{cases} x^2 - (2 + \mu)x, & \text{falls } x \geq 0 \\ (1 - \mu)x, & \text{falls } x < 0. \end{cases}$$

Für $\mu < 1$ ist $\mathcal{L}(x, \mu)$ bzgl. x nach unten unbeschränkt. Für $\mu \geq 1$ wird das Minimum von $\mathcal{L}(x, \mu)$ bzgl. x in $x^* = (2 + \mu)/2$ angenommen.

$$\Rightarrow q(\mu) = \inf_{x \in \mathbb{R}} \mathcal{L}(x, \mu) = \begin{cases} -(2 + \mu)^2/4, & \text{falls } \mu \geq 1, \\ -\infty, & \text{falls } \mu < 1. \end{cases}$$

Der primale Minimalwert ist $\inf(P) = f(1) = -1$, der duale Maximalwert jedoch $\sup(D) = q(1) = -9/4$.

Beachte: Obwohl die primale Zielfunktion hier konvex ist über der zulässigen Menge

$$X = \{x \in \Omega : g(x) \leq 0\} = \{x \in \mathbb{R} : x \geq 0\},$$

tritt eine Dualitätslücke auf. Dies bedeutet, dass alle unteren Schranken $q(\lambda, \mu)$ für den primalen Minimalwert $\inf(P)$ „schlecht“ sind. \diamond

Frage: Wann tritt keine Dualitätslücke auf?

Betrachte die konvexe Optimierungsaufgabe [vgl. (4.4), dort waren jedoch f und g C^1 -Funktionen und $\Omega = \mathbb{R}^n$]

$$\left. \begin{aligned} \text{Minimiere } f(x) \quad & \text{über } x \in \Omega \\ \text{sodass } g_i(x) & \leq 0, \quad i = 1, \dots, k \\ \text{und } Ax & = b. \end{aligned} \right\} \quad (8.6)$$

mit $\Omega \subset \mathbb{R}^n$ konvex sowie $f : \Omega \rightarrow \mathbb{R}$ und $g_i : \Omega \rightarrow \mathbb{R}$ konvex, $i = 1, \dots, k$ und $A \in \mathbb{R}^{(m-k) \times n}$ und $b \in \mathbb{R}^{m-k}$.

Satz 8.7 (Starke Dualität)

Für das konvexe Optimierungsproblem (8.6) gelte $\inf(P) \in \mathbb{R}$, und

- (a) es seien $\Omega = \mathbb{R}^n$ und alle g_i affin-linear, oder
- (b) es existiere ein Vektor $x_0 \in \text{relint}(\Omega)$ mit

$$g_i(x_0) < 0 \quad \text{und} \quad Ax_0 = b.$$

Dann ist das zu (8.6) duale Problem lösbar, und es gilt $\sup(D) = \inf(P)$, d.h., es existiert keine Dualitätslücke.

Dabei ist das **relative Innere** von Ω das Innere bzgl. der Relativtopologie in der affinen Hülle $\text{aff}(\Omega) := \{\sum_{i=1}^k \alpha_i x_i : x_i \in \Omega, \sum_{i=1}^k \alpha_i = 1, k \in \mathbb{N}\}$. Das heißt, $x_0 \in \text{relint}(\Omega)$, falls es eine offene Kugel um x_0 gibt, deren Schnitt mit $\text{aff}(\Omega)$ noch in Ω liegt.

Beachte: Fall (a) verallgemeinert den starken Dualitätssatz für LP !

Beweis: siehe “D.P. Bertsekas: Nonlinear Programming, Athena Scientific, Belmont 1999” für Fall (a) und Satz 6.13, S. 322 in “C. Kanzow, C. Geiger: Theorie und Numerik restringierter Optimierungsaufgaben, Springer, Berlin 2002” für Fall (b) \square

Weitere wichtige Beispiele dualer Probleme sind:

Beispiel 8.8 (QP)

Betrachte das **quadratische Optimierungsproblem** (**quadratisches Programm, QP**, siehe Definition 2.12)

$$\left. \begin{array}{ll} \text{Minimiere} & f(x) = \frac{1}{2}x^\top Qx + c^\top x + \gamma \quad \text{über } x \in \Omega = \mathbb{R}^n \\ \text{sodass} & A_1 x \leq b_1 \\ \text{und} & A_2 x = b_2 \end{array} \right\} \quad (8.7)$$

mit einer symmetrischen positiv semidefiniten Matrix $Q \in \mathbb{R}^{n \times n}$ und $A_1 \in \mathbb{R}^{k \times n}$, $A_2 \in \mathbb{R}^{(m-k) \times n}$, $b_1 \in \mathbb{R}^k$, $b_2 \in \mathbb{R}^{m-k}$.

Beachte: (8.7) ist eine konvexe Optimierungsaufgabe.

Die Lagrange-Funktion

$$\mathcal{L}(x, \lambda, \mu) = \frac{1}{2}x^\top Qx + c^\top x + \gamma + \mu^\top (A_1 x - b_1) + \lambda^\top (A_2 x - b_2).$$

ist konvex bzgl. x . Notwendig und hinreichend für das $\inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda, \mu)$ ist $\mathcal{L}_x(x, \lambda, \mu) = 0$, also

$$Qx + c + A_1^\top \mu + A_2^\top \lambda = 0. \quad (*)$$

Das heißt, es gilt

$$q(\lambda, \mu) = \mathcal{L}(x, \lambda, \mu) \quad \Leftrightarrow \quad (*) \text{ gilt.}$$

Die duale Aufgabe ist demnach gegeben durch

$$\left. \begin{array}{ll} \text{Maximiere} & \mathcal{L}(x, \lambda, \mu) \\ \text{sodass} & Qx + c + A_1^\top \mu + A_2^\top \lambda = 0 \\ \text{und} & \mu \geq 0 \end{array} \right\} \quad \text{über } (x, \lambda, \mu) \in \mathbb{R}^n \times \mathbb{R}^{(m-k)} \times \mathbb{R}^k \quad (8.8)$$

und ist damit wieder ein QP. Ist Q sogar s.p.d., so können wir $(*)$ auflösen und x aus (8.8) eliminieren. \diamond

Beispiel 8.9 (Minimum-Norm-Lösung eines linearen Gleichungssystems)

Wir betrachten

$$\left. \begin{array}{ll} \text{Minimiere} & \|x\| \\ \text{sodass} & Ax = b \end{array} \right\} \quad \text{über } x \in \Omega = \mathbb{R}^n \quad (8.9)$$

mit der euklidischen Norm $\|\cdot\|$, $A \in \mathbb{R}^{k \times n}$ und $b \in \mathbb{R}^k$. Die Lagrange-Funktion ist

$$\mathcal{L}(x, \lambda) = \|x\| + \lambda^\top (Ax - b).$$

Fall (I): Falls $\|A^\top \lambda\| > 1$, dann gibt es ein $z \in \mathbb{R}^n$ mit $\|z\| = 1$ und $z^\top A^\top \lambda > 1$. Wähle $x = -tz$

$$\begin{aligned}\mathcal{L}(x, \lambda) &= t\|z\| - t\lambda^\top Az - \lambda^\top b \quad \text{für } t \geq 0 \\ &= t \left(\underbrace{\|z\| - \lambda^\top Az}_{<0} \right) - \lambda^\top b \\ &\rightarrow -\infty \quad \text{für } t \rightarrow \infty.\end{aligned}$$

Fall (II): Falls $\|A^\top \lambda\| \leq 1$, so gilt

$$\begin{aligned}\mathcal{L}(x, \lambda) &= \|x\| + \lambda^\top Ax - \lambda^\top b \\ &\geq (1 - \|A^\top \lambda\|)\|x\| - \lambda^\top b \quad \text{Cauchy-Schwarz} \\ &\geq -\lambda^\top b,\end{aligned}$$

und dieser Wert wird für $x^* = 0$ angenommen.

Zusammen folgt:

$$q(\lambda) = \begin{cases} -\lambda^\top b, & \text{falls } \|A^\top \lambda\| \leq 1 \\ -\infty, & \text{falls } \|A^\top \lambda\| > 1. \end{cases}$$

[Wegen der Konvexität der Lagrange-Funktion bzgl. x hätten wir auch über den Subgradienten $0 \in \partial_x \mathcal{L}(x, \lambda)$ argumentieren können (siehe § 8).]

Das zu (8.9) duale Problem ist also äquivalent zu:

$$\begin{aligned}\text{Maximiere} \quad & -\lambda^\top b \quad \text{über } \lambda \in \mathbb{R}^k \\ \text{sodass} \quad & \|A^\top \lambda\| \leq 1.\end{aligned} \tag{8.10}$$

Beachte: Das duale Problem hat eine differenzierbare Zielfunktion! \diamond

§ 8.2 Sattelpunkt-Interpretation

Lemma 8.10

Die primale Aufgabe (8.1) ist äquivalent zu

$$\begin{aligned}\text{Minimiere} \quad & \sup_{\substack{\lambda \in \mathbb{R}^{m-k} \\ \mu \in \mathbb{R}^k, \mu \geq 0}} \mathcal{L}(x, \lambda, \mu) \quad \text{über } x \in \Omega.\end{aligned} \tag{8.11}$$

Genauer:

- (a) Ist x^* ein globales (lokales) Optimum von (8.1), so auch von (8.11).
- (b) Ist x^* ein globales (lokales) Optimum von (8.11) (mit endlichem Funktionswert), so auch von (8.1).

Beweis: Es gilt

$$\begin{aligned}\sup_{\lambda; \mu \geq 0} \mathcal{L}(x, \lambda, \mu) &= \sup_{\lambda; \mu \geq 0} [f(x) + \mu^\top g(x) + \lambda^\top h(x)] \\ &= \begin{cases} f(x), & \text{falls } g(x) \leq 0 \text{ und } h(x) = 0 \\ \infty & \text{sonst.} \end{cases}\end{aligned}$$

Damit hat (8.11) dieselben lokalen und globalen Minima mit endlichem Funktionswert wie (8.1). \square

Für den Optimalwert der primalen Aufgabe gilt also

$$\inf(P) = \inf_{x \in \Omega} \sup_{\lambda; \mu \geq 0} \mathcal{L}(x, \lambda, \mu).$$

Aus der Definition 8.1 folgt für die duale Aufgabe (8.3):

$$\sup(D) = \sup_{\lambda; \mu \geq 0} \inf_{x \in \Omega} \mathcal{L}(x, \lambda, \mu),$$

d.h., inf und sup sind vertauscht. Den schwachen Dualitätssatz 8.5 können wir als

$$\sup_{\lambda; \mu \geq 0} \inf_{x \in \Omega} \mathcal{L}(x, \lambda, \mu) \leq \inf_{x \in \Omega} \sup_{\lambda; \mu \geq 0} \mathcal{L}(x, \lambda, \mu) \quad (8.12)$$

schreiben. Starke Dualität bedeutet

$$\sup_{\lambda; \mu \geq 0} \inf_{x \in \Omega} \mathcal{L}(x, \lambda, \mu) = \inf_{x \in \Omega} \sup_{\lambda; \mu \geq 0} \mathcal{L}(x, \lambda, \mu). \quad (8.13)$$

Definition 8.11

Ein Vektor $(x^*, \lambda^*, \mu^*) \in \Omega \times \mathbb{R}^{m-k} \times \mathbb{R}^k$ mit $\mu^* \geq 0$ heißt ein **Sattelpunkt** der Lagrange-Funktion (8.2), wenn gilt:

$$\mathcal{L}(x^*, \lambda, \mu) \leq \mathcal{L}(x^*, \lambda^*, \mu^*) \leq \mathcal{L}(x, \lambda^*, \mu^*) \quad (8.14)$$

für alle $(x, \lambda, \mu) \in \Omega \times \mathbb{R}^{m-k} \times \mathbb{R}^k$ mit $\mu \geq 0$ oder äquivalent:

$$\sup_{\lambda; \mu \geq 0} \mathcal{L}(x^*, \lambda, \mu) = \mathcal{L}(x^*, \lambda^*, \mu^*) = \inf_{x \in \Omega} \mathcal{L}(x, \lambda^*, \mu^*).$$

◇

Beachte: Die rechte Ungleichung besagt, dass x^* ein globales Minimum der Funktion $\mathcal{L}(\cdot, \lambda^*, \mu^*)$ über $x \in \Omega$ ist. Die linke Ungleichung besagt, dass (λ^*, μ^*) ein globales Maximum der Funktion $\mathcal{L}(x^*, \cdot, \cdot)$ über $\lambda \in \mathbb{R}^{m-k}$ und $\mu \in \mathbb{R}^k$ mit $\mu \geq 0$ ist.

Satz 8.12 (Sattelpunkttheorem)

- (a) Es seien $x^* \in \Omega$ und (λ^*, μ^*) mit $\mu^* \geq 0$ globale Optima von (8.1) bzw. (8.3), und es gelte starke Dualität. Dann ist (x^*, λ^*, μ^*) ein Sattelpunkt der Lagrange-Funktion.
- (b) Ist (x^*, λ^*, μ^*) ein Sattelpunkt der Lagrange-Funktion, dann ist x^* ein globales Optimum des primalen Problems (8.1), (λ^*, μ^*) ist ein globales Optimum des dualen Problems (8.3), und es gilt starke Dualität.
- (c) Für differenzierbare konvexe Optimierungsprobleme (4.4) über $\Omega = \mathbb{R}^n$ sind die Sattelpunkte der Lagrange-Funktion genau die KKT-Punkte.

Beweis: (a): Sei $(x, \lambda, \mu) \in \Omega \times \mathbb{R}^{m-k} \times \mathbb{R}^k$ beliebig mit $\mu \geq 0$.

$$\begin{aligned}
\left. \begin{array}{l} \mathcal{L}(x^*, \lambda, \mu) \\ \mathcal{L}(x^*, \lambda^*, \mu^*) \end{array} \right\} &\leq \sup_{\lambda; \mu \geq 0} \mathcal{L}(x^*, \lambda, \mu) \\
&= \inf_{x \in \Omega} \sup_{\lambda; \mu \geq 0} \mathcal{L}(x, \lambda, \mu) \quad \text{wegen primaler Optimalität von } x^* \\
&= \sup_{\lambda; \mu \geq 0} \inf_{x \in \Omega} \mathcal{L}(x, \lambda, \mu) \quad \text{wegen starker Dualität} \\
&= \inf_{x \in \Omega} \mathcal{L}(x, \lambda^*, \mu^*) \quad \text{wegen dualer Optimalität von } (\lambda^*, \mu^*) \\
&\leq \begin{cases} \mathcal{L}(x^*, \lambda^*, \mu^*) \\ \mathcal{L}(x, \lambda^*, \mu^*) \end{cases}
\end{aligned}$$

Aus der oberen und unteren Ungleichung zusammen folgt: (x^*, λ^*, μ^*) ist ein Sattelpunkt. Außerdem ist $\sup_{\lambda; \mu \geq 0} \mathcal{L}(x^*, \lambda, \mu)$ wegen der Optimalität von x^* endlich, also auch alle anderen oben auftretenden Größen.

(b): Sei (x^*, λ^*, μ^*) ein Sattelpunkt.

$$\begin{aligned}
\inf_{x \in \Omega} \sup_{\lambda; \mu \geq 0} \mathcal{L}(x, \lambda, \mu) &\leq \sup_{\lambda; \mu \geq 0} \mathcal{L}(x^*, \lambda, \mu) \quad \text{wegen } x^* \in \Omega \\
&= \mathcal{L}(x^*, \lambda^*, \mu^*) \quad \text{Sattelpunkt-Eigenschaft} \\
&= \inf_{x \in \Omega} \mathcal{L}(x, \lambda^*, \mu^*) \quad \text{Sattelpunkt-Eigenschaft} \\
&\leq \sup_{\lambda; \mu \geq 0} \inf_{x \in \Omega} \mathcal{L}(x, \lambda, \mu) \quad \text{wegen } \mu^* \geq 0.
\end{aligned}$$

Wegen der schwachen Dualität (8.12) gilt überall Gleichheit, d.h.,

$$\begin{aligned}
\inf(P) &= \inf_{x \in \Omega} \sup_{\lambda; \mu \geq 0} \mathcal{L}(x, \lambda, \mu) \\
&= \sup_{\lambda; \mu \geq 0} \mathcal{L}(x^*, \lambda, \mu) \quad (*) \\
&= \mathcal{L}(x^*, \lambda^*, \mu^*) \\
&= \inf_{x \in \Omega} \mathcal{L}(x, \lambda^*, \mu^*) \\
&= \sup_{\lambda; \mu \geq 0} \inf_{x \in \Omega} \mathcal{L}(x, \lambda, \mu) \quad (**) \\
&= \sup(D). \quad (8.15)
\end{aligned}$$

Die Fälle $\inf(P) = \sup(D) = \pm\infty$ können wir ausschließen, denn:

- (a) $\sup_{\lambda; \mu \geq 0} \mathcal{L}(x^*, \lambda, \mu) \neq -\infty$, denn $\{(\lambda, \mu) : \mu \geq 0\} \neq \emptyset$
- (b) $\inf_{x \in \Omega} \mathcal{L}(x, \lambda^*, \mu^*) \neq \infty$, denn $\Omega \neq \emptyset$.

Aus (*) bzw. (**) folgt dann: x^* bzw. (λ^*, μ^*) sind primal bzw. dual optimal, und es gilt starke Dualität.

(c): Sei zunächst (x^*, λ^*, μ^*) ein Sattelpunkt. Dann gilt

$$\sup_{\lambda; \mu \geq 0} \mathcal{L}(x^*, \lambda, \mu) = \mathcal{L}(x^*, \lambda^*, \mu^*) = \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda^*, \mu^*).$$

$$\begin{aligned} \Rightarrow \quad & x^* \text{ ist zulässig (d.h., } g(x^*) \leq 0 \text{ und } h(x^*) = 0) \quad \text{nach Lemma 8.10} \\ \text{und} \quad & \mathcal{L}_x(x^*, \lambda^*, \mu^*) = 0 \quad \text{nach Satz 6.1.} \end{aligned}$$

Wähle $\mu = 0$ und $\lambda = \lambda^*$ in der linken Ungleichung in (8.14).

$$\begin{aligned} \Rightarrow \quad & f(x^*) + (\lambda^*)^\top h(x^*) \leq f(x^*) + (\mu^*)^\top g(x^*) + (\lambda^*)^\top h(x^*) \\ \Rightarrow \quad & (\mu^*)^\top g(x^*) \geq 0 \\ \Rightarrow \quad & (\mu^*)^\top g(x^*) = 0. \end{aligned}$$

Zusammen gilt also die KKT-Bedingung (4.3) für (x^*, λ^*, μ^*) .

Umgekehrt sei (x^*, λ^*, μ^*) ein KKT-Punkt.

$$\begin{aligned} \Rightarrow \quad & \mathcal{L}_x(x^*, \lambda^*, \mu^*) = 0 \\ \Rightarrow \quad & x^* \text{ ist ein globales Minimum der konvexen Funktion } \mathcal{L}(\cdot, \lambda^*, \mu^*), \\ & \text{siehe Satz 2.26} \\ \Rightarrow \quad & \mathcal{L}(x^*, \lambda^*, \mu^*) \leq \mathcal{L}(x, \lambda^*, \mu^*) \quad \text{für alle } x \in \mathbb{R}^n. \end{aligned}$$

Weiter gilt für $(\lambda, \mu) \in \mathbb{R}^{m-k} \times \mathbb{R}^k$ mit $\mu \geq 0$:

$$\begin{aligned} \mathcal{L}(x^*, \lambda, \mu) &= f(x^*) + \underbrace{\mu^\top}_{\geq 0} \underbrace{g(x^*)}_{\leq 0} + \underbrace{\lambda^\top h(x^*)}_{=0} \\ &\leq f(x^*) \\ &= f(x^*) + \underbrace{(\mu^*)^\top g(x^*)}_{=0} + \underbrace{(\lambda^*)^\top h(x^*)}_{=0} \\ &= \mathcal{L}(x^*, \lambda^*, \mu^*), \end{aligned}$$

zusammen also (8.14), d.h., (x^*, λ^*, μ^*) ist ein Sattelpunkt. □

§ 9 Das Subdifferential konvexer Funktionen

Ziel: Verallgemeinerung des Gradienten $\nabla f(\cdot)$ für konvexe (ggf. nicht differenzierbare) Funktionen.

In diesem Abschnitt sei $C \subset \mathbb{R}^n$ konvex und $f : C \rightarrow \mathbb{R}$ eine konvexe Funktion.

Definition 9.1 (Subdifferential)

Es sei $x_0 \in C$.

- (a) Ein Vektor $s \in \mathbb{R}^n$ heißt **Subgradient** von f in x_0 , wenn gilt:

$$f(x) \geq f(x_0) + s^\top (x - x_0) \quad \text{für alle } x \in C. \quad (9.1)$$

- (b) Die Menge $\partial f(x_0)$ aller Subgradienten heißt das **Subdifferential** von f in x_0 . \diamond

Beachte: Die Definition verallgemeinert die Ungleichung (2.6) für konvexe (nicht notwendig differenzierbare) Funktionen. Man sagt: Die rechte Seite in (9.1) ist eine lineare Minorante, die f in x_0 stützt.

Satz 9.2 (Existenz und Eigenschaften des Subdifferentials)

Es sei $x_0 \in \text{int}(C)$. Dann ist $\partial f(x_0) \neq \emptyset$.

Beweis: Betrachte den Epigraphen

$$\text{epi}(f) = \{(x, y) \in \mathbb{R}^n \times \mathbb{R} : y \geq f(x)\}.$$

$\text{epi}(f) \subset \mathbb{R}^n \times \mathbb{R}$ ist konvex (Übungsaufgabe ?). Es ist $(x_0, f(x_0)) \notin \text{int}(\text{epi}(f))$, sonst gäbe es einen Punkt $(x_0, f(x_0) - \varepsilon) \in \text{epi}(f)$, Widerspruch zur Definition von $\text{epi}(f)$.

Nach dem Trennungssatz 2.28 existiert $(s, \alpha) \in \mathbb{R}^n \times \mathbb{R}$, $(s, \alpha) \neq 0$ mit

$$s^\top x + \alpha r \leq s^\top x_0 + \alpha f(x_0) \quad \text{für alle } (x, r) \in \text{epi}(f). \quad (*)$$

Behauptung: Es gilt $\alpha \leq 0$, denn:

Fall (i): Ist $f(x_0) = 0$, wähle $x = x_0$, $r = 1$

$$\Rightarrow s^\top x_0 + \alpha \leq s^\top x_0 \Rightarrow \alpha \leq 0.$$

Fall (ii): Ist $f(x_0) > 0$, wähle $x = x_0$, $r = 2f(x_0)$

$$\Rightarrow s^\top x_0 + 2\alpha f(x_0) \leq s^\top x_0 + \alpha f(x_0) \Rightarrow \alpha \leq 0.$$

Fall (iii): Ist $f(x_0) < 0$, wähle $x = x_0$, $r = 0$

$$\Rightarrow s^\top x_0 \leq s^\top x_0 + \alpha f(x_0) \Rightarrow \alpha \leq 0.$$

Behauptung: Es gilt sogar $\alpha < 0$, denn:

$$x_0 \in \text{int}(C) \Rightarrow x_0 + \varepsilon s \in C \quad \text{für ein } \varepsilon > 0.$$

Setze in (*) $x = x_0 + \varepsilon s$ und $r = f(x)$.

$$\Rightarrow s^\top x + \alpha f(x) = s^\top x_0 + \varepsilon \|s\|^2 + \alpha f(x_0 + \varepsilon s) \leq s^\top x_0 + \alpha f(x_0)$$

$$\Rightarrow \varepsilon \|s\|^2 \leq \alpha [f(x_0) - f(x_0 + \varepsilon s)].$$

Aus $\alpha = 0$ würde $s = 0$ folgen, Widerspruch zu $(s, \alpha) \neq 0$.

Wir können o.B.d.A. $\alpha = -1$ annehmen, dann ergibt (*) mit $r = f(x)$

$$\begin{aligned} s^\top x - f(x) &\leq s^\top x_0 - f(x_0) \quad \text{für alle } x \in C \\ \Rightarrow f(x) &\geq f(x_0) + s^\top(x - x_0) \quad \text{für alle } x \in C, \end{aligned}$$

d.h., $s \in \partial f(x_0)$. □

Frage: Zusammenhang mit Richtungsableitung?

Es sei $x_0 \in \text{int}(C)$ und $d \in \mathbb{R}^n$ eine Richtung. Betrachte den Differenzenquotienten

$$q(t) := \frac{f(x_0 + t d) - f(x_0)}{t}, \quad t > 0. \quad (9.2)$$

Dieser ist definiert zumindest für $t \in (0, \delta)$, sodass $x_0 + t d \in C$ bleibt.

Lemma 9.3 (Monotonie des Differenzenquotienten)

Seien x_0, d und δ wie oben. Dann ist $q(t)$ monoton wachsend auf $(0, \delta)$.

Beweis: Sei $0 < t_1 < t_2 < \delta$. Es gilt $x_0 + t_2 d \in C$ und

$$x_0 + t_1 d = \frac{t_1}{t_2}(x_0 + t_2 d) + \left(1 - \frac{t_1}{t_2}\right)x_0,$$

und C ist konvex, also ist $x_0 + t_1 d \in C$.

$$\begin{aligned} f(x_0 + t_1 d) &= f\left(\frac{t_1}{t_2}(x_0 + t_2 d) + \left(1 - \frac{t_1}{t_2}\right)x_0\right) \leq \frac{t_1}{t_2}f(x_0 + t_2 d) + \left(1 - \frac{t_1}{t_2}\right)f(x_0) \\ \Rightarrow f(x_0 + t_1 d) - f(x_0) &\leq \frac{t_1}{t_2}f(x_0 + t_2 d) - \frac{t_1}{t_2}f(x_0) \\ \Rightarrow q(t_1) &\leq q(t_2). \end{aligned}$$

□

Lemma 9.4

Sei $x_0 \in \text{int}(C)$. Dann gibt es ein $\delta > 0$ mit $U_\delta(x_0) \subset C$, sodass gilt:

- (a) Es existieren $m, M \in \mathbb{R}$ mit $f(x) \in [m, M]$ für alle $x \in U_\delta(x_0)$ [lokale Beschränktheit].
- (b) Es existiert $L(x_0) \geq 0$ mit

$$|f(x) - f(y)| \leq L(x_0) \|x - y\| \quad \text{für alle } x, y \in U_{\delta/2}(x_0)$$

[lokale Lipschitzstetigkeit].

Beweis: (a): Wegen $x_0 \in \text{int}(C)$ existieren Vektoren (Übungsaufgabe ?) $v_0, \dots, v_n \in \mathbb{R}^n$ und $\Delta = \text{co}(\{v_0, \dots, v_n\})$ mit $x_0 \in \text{int}(\Delta)$. Wähle $\delta > 0$ so, dass $U_\delta(x_0) \subset \text{int}(\Delta) \subset C$. Für $x \in U_\delta(x_0)$ gilt $x \in \Delta$, also

$$\begin{aligned} x &= \sum_{i=0}^n \alpha_i v^i, \quad \text{mit } \alpha_i \geq 0 \text{ und } \sum_{i=0}^n \alpha_i = 1 \\ \Rightarrow f(x) &= f\left(\sum_{i=0}^n \alpha_i v^i\right) \leq \sum_{i=0}^n \alpha_i f(v^i) \leq \underbrace{\max_{i=0, \dots, n} f(v^i)}_{=: M} \sum_{i=0}^n \alpha_i = M. \end{aligned}$$

Sei nun $s \in \partial f(x_0)$. Wegen $f(x) \geq f(x_0) + s^\top(x - x_0)$ ist f auf $U_\delta(x_0)$ nach unten beschränkt durch $m := f(x_0) - \delta \|s\|$.

(b): Es seien $x, y \in U_{\delta/2}(x_0)$ und $x \neq y$. Setze $v := y + \frac{\delta}{2} \frac{y-x}{\|y-x\|}$. Dann ist $v \in U_{\delta}(x_0)$, denn

$$\|v - x_0\| \leq \|v - y\| + \|y - x_0\| < \frac{\delta}{2} + \frac{\delta}{2} = \delta.$$

Setze $\alpha := \frac{2\|y-x\|}{\delta+2\|y-x\|} \in (0, 1)$.

$$\begin{aligned} f(y) - f(x) &= f(\alpha v + (1 - \alpha)x) - f(x) && \text{nachrechnen} \\ &\leq \alpha f(v) + (1 - \alpha)f(x) - f(x) && \text{Konvexität von } f \\ &= \alpha(f(v) - f(x)) \\ &\leq \alpha(M - m) && \text{Teil (a)} \\ &\leq \frac{2\|y-x\|}{\delta}(M - m) && \text{denn } \alpha \leq \frac{2\|y-x\|}{\delta}. \end{aligned}$$

Durch Vertauschen von x und y bekommt man die Behauptung mit $L(x_0) := \frac{2(M-m)}{\delta}$. \square

Wegen Lemma 9.4 gilt für den Differenzenquotienten aus (9.2)

$$|q(t)| = \frac{|f(x_0 + t d) - f(x_0)|}{t} \leq L(x_0) \frac{\|x_0 + t d - x_0\|}{t} = L(x_0) \|d\|$$

für alle $t \in (0, \delta)$ mit einem $\delta > 0$. $q(t)$ ist also auf $(0, \delta)$ monoton wachsend und beschränkt, d.h., es existiert

$$\lim_{t \searrow 0} q(t) = \inf_{t > 0} q(t). \quad (9.3)$$

Definition 9.5 (Richtungsableitung)

Sei $x_0 \in \text{int}(C)$. Der Ausdruck

$$\delta f(x_0; d) := \lim_{t \searrow 0} q(t) = \lim_{t \searrow 0} \frac{f(x_0 + t d) - f(x_0)}{t} \quad (9.4)$$

heißt **Richtungsableitung** von f im Punkt x_0 in Richtung $d \in \mathbb{R}^n$. \diamond

Beachte: Konvexe Funktionen sind im Inneren ihres Definitionsbereiches richtungsdifferenzierbar.

Bemerkung 9.6 (Eigenschaften der Richtungsableitung)

Die Richtungsableitung $\delta f(x_0; \cdot)$ aus (9.4) hat folgende Eigenschaften:

(a) positiv homogen:

$$\delta f(x_0; \lambda d) = \lambda \delta f(x_0; d) \quad \text{für alle } \lambda > 0$$

(b) subadditiv:

$$\delta f(x_0; d_1 + d_2) \leq \delta f(x_0; d_1) + \delta f(x_0; d_2)$$

(c) konvex:

$$\delta f(x_0; \lambda d_1 + (1 - \lambda) d_2) \leq \lambda \delta f(x_0; d_1) + (1 - \lambda) \delta f(x_0; d_2)$$

für alle $\lambda \in [0, 1]$

(d) beschränkt durch die lokale Lipschitz-Konstante aus Lemma 9.4:

$$|\delta f(x_0; d)| \leq L(x_0) \|d\| \quad \text{für alle } d \in \mathbb{R}^n;$$

insbesondere gilt $\delta f(x_0; 0) = 0$

(e) Lipschitzstetig bzgl. d :

$$|\delta f(x_0; d_1) - \delta f(x_0; d_2)| \leq L(x_0) \|d_1 - d_2\| \quad \text{für alle } d_1, d_2 \in \mathbb{R}^n. \quad \diamond$$

Beweis: eigene Übung

□

Satz 9.7 (Charakterisierung des Subdiff. durch Richtungsableitung)

Sei $x_0 \in \text{int}(C)$. Dann gilt

(a)

$$\partial f(x_0) = \{s \in \mathbb{R}^n : s^\top d \leq \delta f(x_0; d) \text{ für alle } d \in \mathbb{R}^n\}.$$

[Subgradienten liegen „unterhalb“ der Richtungsableitung]

(b)

$$\delta f(x_0; d) = \max_{s \in \partial f(x_0)} s^\top d \quad \text{für alle } d \in \mathbb{R}^n.$$

Beweis: (a) „ \subset “: Sei $s \in \partial f(x_0)$ und $d \in \mathbb{R}^n$. Sei $\delta > 0$ so, dass $x_0 + t d \in C$ für $t \in (0, \delta)$.

$$\Rightarrow f(x_0 + t d) \geq f(x_0) + t s^\top d \quad \text{für alle } t \in (0, \delta)$$

$$\Rightarrow \frac{f(x_0 + t d) - f(x_0)}{t} \geq s^\top d \quad \text{für alle } t \in (0, \delta)$$

$$\Rightarrow \delta f(x_0; d) \geq s^\top d$$

(a) „ \supset “: Sei $s \in \mathbb{R}^n$ mit $s^\top d \leq \delta f(x_0; d)$ für alle $d \in \mathbb{R}^n$. Sei $x \in C$ beliebig, setze $d := x - x_0$.

$$\Rightarrow s^\top d \leq \delta f(x_0; d) \leq \frac{f(x_0 + t d) - f(x_0)}{t} \quad \text{für alle } t \in (0, 1] \quad \text{wegen (9.3)}$$

$$\Rightarrow f(x) = f(x_0 + d) \geq f(x_0) + s^\top d = f(x_0) + s^\top (x - x_0) \quad \text{mit } t = 1,$$

d.h., $s \in \partial f(x_0)$.

(b): Sei $d \in \mathbb{R}^n$. Nach Teil (a) gilt

$$\begin{aligned} \delta f(x_0; d) &\geq s^\top d \quad \text{für alle } s \in \partial f(x_0) \\ \Rightarrow \delta f(x_0; d) &\geq \sup_{s \in \partial f(x_0)} s^\top d. \end{aligned} \quad (*)$$

Nach Bemerkung 9.6 (c) ist $\delta f(x_0; \cdot)$ konvex auf \mathbb{R}^n . Nach Satz 9.2 existiert also ein Subgradient $s_d \in \mathbb{R}^n$ an $\delta f(x_0; \cdot)$ im Punkt d , d.h.,

$$\delta f(x_0; y) \geq \delta f(x_0; d) + s_d^\top (y - d) \quad \text{für alle } y \in \mathbb{R}^n \quad (**)$$

$$y = 0 \quad \Rightarrow \quad 0 \geq \delta f(x_0; d) - s_d^\top d$$

$$\Rightarrow s_d^\top d \geq \delta f(x_0; d)$$

$$y = 2d \quad \Rightarrow \quad 2\delta f(x_0; d) = \delta f(x_0; 2d) \geq \delta f(x_0; d) + s_d^\top (2d - d)$$

$$\Rightarrow \delta f(x_0; d) \geq s_d^\top d$$

Also gilt $s_d^\top d = \delta f(x_0; d)$, und aus $(**)$ folgt

$$\delta f(x_0; y) \geq \delta f(x_0; d) + s_d^\top (y - d) = s_d^\top y \quad \text{für alle } y \in \mathbb{R}^n,$$

d.h., $s_d \in \partial f(x_0)$ nach Teil (a). Aus (*) folgt jetzt

$$\delta f(x_0; d) \geq \sup_{s \in \partial f(x_0)} s^\top d \geq s_d^\top d = \delta f(x_0; d),$$

und das Supremum wird durch s_d angenommen, was (b) beweist. \square

Korollar 9.8 (Kompaktheit des Subdifferentials)

Sei $x_0 \in \text{int}(C)$. Dann ist $\partial f(x_0) \neq \emptyset$, konvex und kompakt.

Beweis: $\partial f(x_0) \neq \emptyset$ wegen Satz 9.2. Die Konvexität und Abgeschlossenheit von $\partial f(x_0)$ folgen direkt aus der Definition 9.1. Für $s \in \partial f(x_0)$ gilt nach Satz 9.7 (a)

$$s^\top d \leq \delta f(x_0; d) \leq L(x_0) \|d\| \quad \text{für alle } d \in \mathbb{R}^n$$

mit der lokalen Lipschitz-Konstante aus Bemerkung 9.6 (c). Mit $d = s$ folgt $\|s\| \leq L(x_0)$, d.h., $\partial f(x_0)$ ist beschränkt. \square

Satz 9.9 (Subdifferential einer differenzierbaren Funktion)

Sei $x_0 \in \text{int}(C)$ und f differenzierbar in x_0 . Dann gilt

$$(a) \quad \delta f(x_0; d) = \nabla f(x_0)^\top d \quad \text{für alle } d \in \mathbb{R}^n$$

$$(b) \quad \partial f(x_0) = \{\nabla f(x_0)\}.$$

Beweis:

(a): Aus der Differenzierbarkeit von f in x_0 folgt

$$\lim_{\|h\| \rightarrow 0} \frac{f(x_0 + h) - f(x_0) - \nabla f(x_0)^\top h}{\|h\|} = 0.$$

Setze $h := t d$ mit $t \searrow 0$ und $d \in \mathbb{R}^n$.

(b): „ \supset “: Es gilt $\nabla f(x_0) \in \partial f(x_0)$, denn nach Satz 2.24 (a) [stetige Differenzierbarkeit wird nicht benötigt] ist

$$f(x) - f(x_0) \geq \nabla f(x_0)^\top (x - x_0)$$

für alle $x \in C$.

„ \subset “: Nach Satz 9.7 gilt

$$\partial f(x_0) = \{s \in \mathbb{R}^n : s^\top d \leq \nabla f(x_0)^\top d \text{ für alle } d \in \mathbb{R}^n\}.$$

Für $s \in \partial f(x_0)$ gilt also $(s - \nabla f(x_0))^\top d \leq 0$ für alle $d \in \mathbb{R}^n$, d.h., $s = \nabla f(x_0)$. \square