Für Funktionen $f: \mathbb{R} \supset I \to \mathbb{R}$ haben wir die folgenden hinreichenden Kriterien für lokale Extrema gelernt:

$f'(x_0) = 0$ und $f''(x_0) > 0 \Rightarrow f$ besitzt in x_0 ein lokales Minimum

und, äquivalent dazu

$f'(x_0) = 0$ und $f''(x_0) < 0 \Rightarrow f$ besitzt in x_0 ein lokales Maximum.

Im letzten Abschnitt haben wir bereits gesehen, dass die notwendige Bedingung $f'(x_0) = 0$ für Funktionen mehrerer Veränderlicher zu ersetzen ist durch die Forderung $\nabla f(x_0) = 0$,

NOTWENDIG FÜR LOK. EXTREMA (ohne Randextrema)

also $f'(x_0) = 0 \rightsquigarrow \nabla f(x_0) = 0$

$f'(x_0) = 0$: hier ist eine einzelne Gleichg zu lösen.

$\nabla f(x_0) = 0$: i. allg. nichtlineares GLS. mit n Gleichungen für n Unbekannte zu lösen.

Was tritt an die Stelle der hinreichenden Bedingungen

$f''(x_0) > 0$ (für ein Min.) bzw. $f''(x_0) < 0$ (für ein Max.)?

Dazu bildet man aus den zweiten partiellen Ableitungen $\frac{\partial^2 f}{\partial x_k \partial x_j}(x)$ eine quadratische $n \times n$-Matrix, die sogenannte Hesse-Matrix

Def.: Es sei $f: \mathbb{R}^n \supset \Omega \to \mathbb{R}$ in $x_0 \in \Omega$ zweimal partiell nach allen Variablen $x_1, \dots, x_n$ differenzierbar. Dann heißt

$$\operatorname{Hess} f(x_0) := \left(\frac{\partial^2 f}{\partial x_j \partial x_k}(x_0) \right)_{1 \le j,k \le n}$$

die <u>Hesse-Matrix</u> von f in x_0.

Bem.: (1) Ausgeschrieben:

$$\left(\frac{\partial^2 f}{\partial x_j \partial x_k}(x_0) \right)_{1 \le j,k \le n} = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x_0) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x_0) \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x_0) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(x_0) \end{pmatrix}$$

(2) Sind alle 2. partiellen Ableitungen stetig, so gilt nach dem Satz v. Schwarz, daß

$$\frac{\partial^2 f}{\partial x_k \partial x_j}(x_0) = \frac{\partial^2 f}{\partial x_j \partial x_k}(x_0).$$

In diesem Fall ist die Hessematrix von f also eine symmetrische Matrix. Dabei heißt eine Matrix $A = (a_{ij})_{1 \le i,j \le n}$ symmetrisch, wenn $a_{ij} = a_{ji}$ für alle $i,j \in \{1, \dots, n\}$.

Wie kann man anhand der Hesse-Matrix feststellen, ob in einer kritischen Stelle x_0 von f ein Minimum, Maximum oder ein Sattelpunkt vorliegt? Dazu müssen wir den folgenden Begriff einführen:

Def.: Eine $n\times n$-Matrix $A = (a_{ij})_{1\le i,j\le n}$ heißt

(a) positiv semidefinit, wenn $\langle x, Ax\rangle \ge 0$ für alle $x\in\mathbb{R}^n$,

(b) positiv definit, wenn $\langle x, Ax\rangle > 0$ für alle $x\in\mathbb{R}^n\setminus\{0\}$,

(c) negativ semidefinit, wenn $-A$ positiv semidefinit ist,

(d) negativ definit, wenn $-A$ positiv definit ist,

(e) indefinit, wenn $x_1\in\mathbb{R}^n$ und $x_2\in\mathbb{R}^n$ existieren, so daß $\langle x_2, Ax_2\rangle < 0 < \langle x_1, Ax_1\rangle$ ist.

Bem. + Bsp.: (1) Für $x = (x_1,\dots,x_n)^T$ und $A = (a_{ij})_{1\le i,j\le n}$ ist $\langle x, Ax\rangle = \sum_{i,j=1}^{n} x_i a_{ij} x_j$.

(2) Ist $A = \begin{pmatrix} a & 0 \\ 0 & b\end{pmatrix}$, eine 2×2-Matrix in Diagonalsgestalt, $x = \begin{pmatrix} x_1 \\ x_2\end{pmatrix}$, so haben wir $\langle x, Ax\rangle = ax_1^2 + bx_2^2$. Daran lesen wir ab: A ist

positiv (semi-) definit, wenn $a, b > 0$ (bzw. $a, b \ge 0$),

negativ (semi-) definit, wenn $a, b < 0$ (bzw. $a, b \le 0$),

indefinit, wenn $a < 0 < b$ oder $a > 0 > b$.

Welche Rolle spielt die Definitheit der Hesse-Matrix einer Funktion f für die Bestimmung einer lokalen Extrema? Diesbezüglich gilt der folgende

Satz: Es seien $\Omega \subset \mathbb{R}^n$ offen, $x_0 \in \Omega$, $f : \Omega \to \mathbb{R}$ zweimal stetig partiell diff'bar und $\nabla f(x_0) = 0$. Dann gilt:

(1) Ist $\operatorname{Hess} f(x_0)$ negativ definit, so besitzt f in x_0 ein isoliertes lokales Maximum;

(2) ist $\operatorname{Hess} f(x_0)$ positiv definit, so besitzt f in x_0 ein isoliertes lokales Minimum;

(3) ist $\operatorname{Hess} f(x_0)$ indefinit, so besitzt f in x_0 kein Extremum.

Begründung: Verallgemeinerung des MWS
(für Aussage (2))

$$f(x_0+h) = f(x_0) + \langle \nabla f(\xi), h \rangle \quad (\xi \in [x_0, x_0+h])$$

zeigt man: $f(x_0+h) = f(x_0) + \langle \nabla f(x_0), h \rangle + \frac{1}{2} \langle h, \operatorname{Hess} f(\xi) \cdot h \rangle$.

Ist x_0 eine kritische Stelle von f, also $\nabla f(x_0) = 0$, so reduziert sich dies auf

$$f(x_0+h) = f(x_0) + \frac{1}{2} \langle h, \operatorname{Hess} f(\xi) h \rangle .$$

Ist $\operatorname{Hess} f(x_0)$ positiv def., so gilt dies aus Stetigkeitsgründen auch für $\operatorname{Hess} f(\xi)$, jedenfalls, wenn $|h|$ klein genug ist. Also ist der zweite Summand rechts dann positiv und damit

$$f(x_0+h) > f(x_0),$$

sofern h in einer hinreichend kleinen Nullumgebung liegt.

Um den Satz anzuwenden, müssen wir entscheiden können, ob eine Matrix pos./neg. definit / semidefinit oder indefinit ist. Dazu gibt es im wesentlichen zwei Kriterien!

1. Eigenwertkriterium

(Erinnerung: $\lambda \in \mathbb{C}$ heißt ein Eigenwert von $A \in \mathbb{R}^{n \times n}$, falls $x \in \mathbb{R}^n \setminus \{0\}$ (ein sog. Eigenvektor) existiert, für den $Ax = \lambda x$ gilt.)

Für symmetrische Matrizen gilt das folgende

<u>Definitheitskrit.</u>: Eine symmetrische, reelle $n \times n$-Matrix A mit Eigenwerten $\lambda_1, \dots, \lambda_n$ ist

(a) positiv semidefinit $\Leftrightarrow \lambda_i \geq 0 \ \forall i \in \{1, \dots, n\}$

(b) positiv definit $\Leftrightarrow \lambda_i > 0 \ \forall i \in \{1, \dots, n\}$

(c) negativ semidefinit $\Leftrightarrow \lambda_i \leq 0 \ \forall i \in \{1, \dots, n\}$

(d) negativ definit $\Leftrightarrow \lambda_i < 0 \ \forall i \in \{1, \dots, n\}$

(e) indefinit $\Leftrightarrow \exists \ i, j \in \{1, \dots, n\}$, so dass $\lambda_i < 0 < \lambda_j$.

Der Fall $n = 2$: Eine symmetrische 2×2 Matrix hat die allgemeine Form: $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ mit $a, b, c \in \mathbb{R}$.

Die Eigenwerte sind allgemein die Nullstellen des charakteristischen Polynoms $P(\lambda) = \det(\lambda I - A)$, im Fall einer symmetrischen 2×2-Matrix also von

$$P(\lambda) = \det \begin{pmatrix} a-\lambda & b \\ b & c-\lambda \end{pmatrix} = (a-\lambda)(c-\lambda) - b^2$$

$$= \lambda^2 - (a+c)\lambda + ac - b^2.$$

$$P(\lambda) = 0 \Leftrightarrow \lambda_{1,2} = \frac{a+c}{2} \pm \sqrt{\frac{(a+c)^2}{4} + ac + b^2}$$

In diesem Fall gilt:

A ist (semi-) definit $\Leftrightarrow \lambda_1 \lambda_2 > 0 \quad (\lambda_1 \lambda_2 \geq 0)$

$\Leftrightarrow \quad ac - b^2 > 0 \quad (ac - b^2 \geq 0)$ (d.h. auch: A ist indefinit, genau dann, wenn $ac - b^2 < 0$!)

und zwar positiv (semi-) definit,

genau dann, wenn $\lambda_1 + \lambda_2 = a + c > 0$ (bzw. $a+c \geq 0$).

11.07.

Für große n kann die Berechnung von Eigenwerten problematisch sein (Rechenaufwand, keine exakten Lösungen!).

In diesem Fall hilft das

2. Determinantenkriterium (von Hurwitz):

Für eine symmetrische Matrix $A = (a_{ij})_{1 \leq i,j \leq n}$ bildet man die Untermatrizen $A_k = (a_{ij})_{1 \leq i,j \leq k}$, wobei $k \leq n$, also

$$A_1 = (a_{11}),\ A_2 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},\ A_3 = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \dots$$

Dann gilt: A ist positiv definit $\Leftrightarrow \det(A_k) > 0$ für alle $k \in \{1, \dots, n\}$.

<u>Vorsicht</u>: Will man hiermit nachweisen, dass A negativ definit ist, muß man das Krit. auf $-A$ anwenden.

Bsp. Für $(x,y) \in \mathbb{R}^2$ sei

$$f(x,y) = x^2 + y^2 - \lambda xy + 2(x-y)$$ mit einem Parameter $\lambda > 0$

1. Berechnung der ersten partiellen Ableitungen:

$$\frac{\partial f}{\partial x}(x,y) = 2x - \lambda y + 2, \quad \frac{\partial f}{\partial y}(x,y) = 2y - \lambda x - 2$$

Hieraus bilden wir den Gradienten

$$\nabla f(x,y) = (2x - \lambda y + 2, 2y - \lambda x - 2)$$

2. Bestimmung der kritischen Stellen:

Notwendige Bedingung: $\nabla f(x,y) = 0 \; (=(0,0))$. Das ist hier ein Gleichungssystem aus zwei Gleichungen:

(i) $0 = \frac{\partial f}{\partial x}(x,y) = 2x - \lambda y + 2$

(ii) $0 = \frac{\partial f}{\partial y}(x,y) = 2y - \lambda x - 2$

In diesem speziellen Fall handelt es sich sogar um ein lineares GLS, was wir systematisch lösen können. Ad hoc kommen wir jedoch recht schnell zu einer Lösung, indem wir Summe und Differenz beider Gleichungen bilden (das ist auch für nichtlineare Gleichungssysteme möglich, aber nicht immer zielführend).

(i) + (ii) $\quad 0 = (2-\lambda)(x+y)$

(i) − (ii) $\quad 0 = (2+\lambda)(x-y) \cancel{+} + 4$

Nun sind offenbar zwei Fälle zu unterscheiden:

(a) $\lambda = 2 \Rightarrow$ 1. Gleichung ist stets erfüllt. Aus der zweiten ergibt sich: $0 = 4(x-y)+4 \Rightarrow y = x+1$.

Die Menge der kritischen Punkte ist hier also die Gerade

$\{(x,y)\in\mathbb{R}^2 : y = x+1\}$.

(b) $\lambda \neq 2$: Aus der ersten Gleichung folgt dann $x = -y$.

Setzen wir das in die zweite ein, erhalten wir

$2(2+\lambda)x = -4 \Rightarrow x = \frac{-2}{2+\lambda}$. Es gibt also nur einen kritischen Punkt, nämlich $(x,y) = \frac{2}{2+\lambda}(-1,1)$

3. Berechnung der zweiten partiellen Ableitungen:

$$\frac{\partial^2 f}{\partial x^2}(x,y) = \frac{\partial}{\partial x}\frac{\partial f}{\partial x}(x,y) = \frac{\partial}{\partial x}(2x-\lambda y+2) = 2$$

$$\frac{\partial^2 f}{\partial x \partial y}(x,y) = \frac{\partial^2 f}{\partial y \partial x}(x,y) = \frac{\partial}{\partial y}(2x-\lambda y+2) = -\lambda$$

Satz v. Schwarz

$\longrightarrow$ allgemeiner Hinweis!

$$\frac{\partial^2 f}{\partial y^2}(x,y) = \frac{\partial}{\partial y}(2y-\lambda x-2) = 2$$

4. Anordnung zur Hesse-Matrix:

$$\operatorname{Hess} f(x,y) = \begin{pmatrix} 2 & -\lambda \\ -\lambda & 2 \end{pmatrix}$$

Hier: Unabhängig von x und y. Das liegt daran, daß wir hier ein quadratisches Polynom untersuchen. Im allgemeinen ist $\operatorname{Hess} f(x,y)$ abhängig von x und y, und es kommt auf die Definitheit in den kritischen Stellen an!

5. Feststellung der Definitheit der Hesse-Matrix, hier: in Abhängigkeit vom Parameter $\lambda > 0$.

$$\det \operatorname{Hess} f(x,y) = \begin{vmatrix} 2 & -\lambda \\ -\lambda & 2 \end{vmatrix} = 4-\lambda^2 \begin{cases} >0 & \text{für } 0<\lambda<2 \\ =0 & \text{"} \quad \lambda = 2 \\ <0 & \text{"} \quad \lambda > 2 \end{cases}$$

Fallunterscheidung:

(α) $0<\lambda<2$: Die Hesse-Matrix ist definit, und zwar positiv, weil $\underbrace{a+c}_{\text{Summe der Diagonalelemente}} = 2+2=4>0$.

Das bedeutet: In der kritischen Stelle $(x_0,y_0)=\frac{2}{2+\lambda}(-1,1)$ liegt ein (isoliertes) lokales Minimum vor.

(β) $\lambda=2$: Hier ist $\det \operatorname{Hess} f(x,y)=0$, das heißt $\lambda_0=0$ ist ein Eigenwert der Matrix, sie ist weder definit noch indefinit, aber semidefinit. Unsere Kriterien erlauben keine Aussage. Aber:

$$f(x,y) = x^2+y^2-2xy+2(x-y) = (x-y)^2+2(x-y)+1-1$$
$$= (x-y+1)^2-1 \geq -1$$

Auf der Geraden $\{y=x+1\}$ wird der erste Summand des vorletzten Ausdrucks tatsächlich Null. In den kritischen Stellen haben wir also ein globales, nicht-isoliertes Minimum.

(γ) $\lambda>2$: In diesem Fall ist $\operatorname{Hess} f(x,y)$ indefinit, es liegt in der kritischen Stelle ein Sattelpunkt vor. Für $\lambda>2$ hat f also keine Extrema.

Als nächstes soll eine Funktion untersucht werden, die tatsächlich von n Veränderlichen abhängt.

Bsp.: Gegeben seien $a^{(1)}, \dots, a^{(N)} \in \mathbb{R}^n$ (feste Punkte im n-dimensionalen Raum). Gesucht ist nach einem weiteren Punkt $x_0 \in \mathbb{R}^n$, für den die Summe

$$\sum_{k=1}^{N} |x - a^{(k)}|^2 = \sum_{k=1}^{N} \sum_{i=1}^{n} (x_i - a_i^{(k)})^2 =: f(x)$$

der Abstandsquadrate minimal wird.

Die notwendige Bedingung $\nabla f(x_0) = 0$ besteht aus den n Gleichungen

$$0 = \frac{\partial f}{\partial x_j}(x) = \sum_{k=1}^{N} \sum_{i=1}^{n} \frac{\partial}{\partial x_j} (x_i - a_i^{(k)})^2$$

$$= 2 \sum_{k=1}^{N} x_j - a_j^{(k)} = 2\Big(N x_j - \sum_{k=1}^{N} a_j^{(k)}\Big)$$

Dies läßt sich zusammenfassen zu (für die Minimalstelle x_0): $x_0 = \frac{1}{N} \sum_{k=1}^{N} a^{(k)}$,

was man physikalisch als den Schwerpunkt des Systems aus den Massepunkten $a_1, \dots, a_N$ auffassen würde.

Die Hesse-Matrix wird gebildet aus den zweiten Ableitungen

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \frac{\partial}{\partial x_i} \Big(2N x_j - \sum_{k=1}^{N} a_j^{(k)}\Big) = 2N \cdot \delta_{ij},$$

also ist die Hesse-Matrix gegeben durch

$$\operatorname{Hess} f(x) = 2N \cdot E_n \leftarrow (n\times n\text{-Einheitsmatrix}).$$

Diese ist positiv definit (Eigenwerte oder Hurwitz oder Definition), also liegt ein isoliertes lokales Minimum in $x_0 = \frac{1}{N}\sum_{k=1}^{N} a^{(k)}$ vor. Dies ist zugleich ein globales, denn wir haben

$$\lim_{x\to\infty} f(x) = \lim_{x\to\infty} \sum_{k=1}^{N} |x - a^{(k)}|^2 = \infty$$

(in jede Richtung), da f stetig ist, muß ein globales Min. existieren.

Als letztes Beispiel betrachten wir

Bsp. Gewinnfunktionen $G : (0,\infty)^n \to \mathbb{R}$

allgemeine Struktur: $G(x) = \underbrace{\underbrace{p(x)}_{\text{Preis}} \cdot \underbrace{f(x)}_{\text{Produktionsfunktion}}}_{\text{Erlösfunktion}} - \underbrace{K(x)}_{\text{Kostenfunktion}}$

dabei ist $x = (x_1, \ldots, x_n) \in (0,\infty)^n$,

x_i = Komponenten zur Herstellung eines Produkts

Vereinfachende Annahmen:

$p(x) = p$ (konstanter Preis)

$K(x) = \sum_{i=1}^{n} k_i x_i$ (lineare Kostenfunktion, $k_i > 0$)

(später) f eine Cobb-Douglas-Funktion

Damit lauten also die notwendigen Bedingungen zur Gewinnmaximierung:

$$\frac{\partial G}{\partial x_j}(x) = p \cdot \frac{\partial f}{\partial x_j}(x) - k_j \stackrel{!}{=} 0$$

Nun sei also $f(x) = c \cdot \prod_{i=1}^{n} x_i^{s_i}$, wobei wir zur weiteren Vereinfachung $p \cdot c = 1$ annehmen. $\frac{\partial f}{\partial x_j}(x) = \frac{s_j}{x_j} \cdot f(x)$ wurde bereits berechnet, so dass wir das GlS.

$$\frac{s_j}{x_j} f(x) \stackrel{!}{=} k_j \quad \text{bzw.} \quad x_j \stackrel{!}{=} \frac{s_j}{k_j} \cdot f(x)$$

zu lösen haben. Jede Lösung $x = (x_1, \dots, x_n)$ dieses Systems muß also ein Vielfaches von $\left(\frac{s_1}{k_1}, \dots, \frac{s_n}{k_n}\right)$ sein, also gehen wir mit dem

$$\text{Ansatz:} \qquad x = \lambda\left(\frac{s_1}{k_1}, \dots, \frac{s_n}{k_n}\right)$$

in das GlS. ein und erhalten für die j-te Komponente

$$\lambda \frac{s_j}{k_j} = f\left(\lambda\left(\frac{s_1}{k_1}, \dots, \frac{s_n}{k_n}\right)\right) \cdot \frac{s_j}{k_j} \qquad (*)$$

Beachten wir $f(x) = \prod_{i=1}^{n} x_i^{s_i}$ und setzen $s = \sum_{i=1}^{n} s_i$

$$\lambda^{1-s} = \prod_{i=1}^{n} \left(\frac{s_i}{k_i}\right)^{s_i} \quad \text{bzw.} \quad \lambda = \prod_{i=1}^{n} \left(\frac{s_i}{k_i}\right)^{\frac{s_i}{1-s}}$$

Wir erhalten also nur eine kritische Stelle, nämlich

$$x_0 = \left(\prod_{i=1}^{n} \left(\frac{s_i}{k_i}\right)^{\frac{s_i}{1-s}}\right)\left(\frac{s_1}{k_1}, \dots, \frac{s_n}{k_n}\right)$$

Für $0 < s < 1$ ist ein Maximum zu erwarten, denn $\lim_{x \to \infty} G(x) = -\infty$ (nicht jedoch für $s > 1$, da dann $\lim_{x \to \infty} G(x) = \infty$ ist.)

Bei Anwendung des hinreichenden Kriteriums an die Hesse-Matrix erweist sich hier als schwierig. Für die zweiten partiellen Ableitungen der Cobb-Douglas-Funktion haben wir bereits berechnet:

$$\frac{\partial^2 G}{\partial x_j \partial x_k}(x) = \frac{\partial^2 f}{\partial x_j \partial x_j}(x) = f(x)\left(\frac{S_j S_k}{x_j x_k} - \delta_{jk}\frac{S_k}{x_k^2}\right)$$

und in der kritischen Stelle $x_0 = \lambda\left(\frac{S_1}{K_1}, \dots, \frac{S_n}{K_n}\right)$

$$\frac{\partial^2 G}{\partial x_j \partial x_k}(x) \underset{(*)}{=} \underbrace{\lambda^{-1}}_{>0}\left(k_j k_k \left(1 - \frac{\delta_{jk}}{S_k}\right)\right)$$

Damit erhalten wir für die Hesse-Matrix

$$\text{Hess } G(x_0) = -\lambda^{-1}\begin{pmatrix} k_1^2\left(\frac{1}{S_1}-1\right) & -k_1k_2 & \cdots & -k_1k_n \\ -k_1k_2 & k_2^2\left(\frac{1}{S_2}-1\right) & & \vdots \\ \vdots & & \ddots & \vdots \\ -k_1k_n & \cdots & \cdots & k_n^2\left(\frac{1}{S_n}-1\right) \end{pmatrix}$$

Mit Hilfe der Multilinearität der Determinante erhalten wir weiter

$$|-\text{Hess } G(x_0)| = \lambda^{-n}\prod_{j=1}^{n} k_j^2 \begin{vmatrix} \frac{1}{S_1}-1 & -1 & \cdots & -1 \\ -1 & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ -1 & \cdots & \cdots & \frac{1}{S_n}-1 \end{vmatrix}$$

und die zuletzt angeschriebene Determinante erweist sich tatsächlich als positiv, wenn $S = \sum_{i=1}^{n} S_i < 1$ ist.

(Bsp.: $n=2$: $\frac{(1-S_1)(1-S_2)}{S_1 \quad S_2} - \frac{S_1S_2}{S_1S_2} > 0 \iff 1 - S_1 - S_2 > 0$

$\iff S_1 + S_2 < 1$)

Insofern liegt nach Hurwitz tatsächlich ein isoliertes Maximum in x_0 vor, wie wir bereits vermutet hatten.