Erklärung zu Extrema unter NB

Extrema bei Funktionen mit mehreren Variablen mit Nebenbedingungen:

Lösungsverfahren: **(1) Variablensubstitution**

(2) Lagrangeansatz

⇨ Das Optimierungsverfahren mit Lagrange-Multiplikatoren

Lagrange-Verfahren

Idee nach einem von Joseph-Louis Lagrange entwickelten:
Man erweitert die Zielfunktion f, indem man jede Nebenbedingung mit einem sogenannten Lagrange-Multiplikator als zusätzliche Variable multipliziert und diesen Term dann additiv mit der Zielfunktion verknüpft.
Danach wird per Differentialrechnung und mittels Lösung eines Gleichungssystems das Austauschverhältnis der am Prozess beteiligten Faktoren, das Maximum/Minimum der Zielfunktion und die Werte des Lagrangemultiplikators bzw. der Lagrangemultiplikatoren berechnet, um diese dann für eine Sensitivitätsanalyse verwenden zu können.

Lagrangeansatz für den Fall n=2 und eine Nebenbedingung:

$$L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda \cdot g(x_1, x_2)$$

$f(x_1, x_2)$ ist die Zielfunktion $\quad g(x_1, x_2)$ ist die Nebenbedingung

Vorgehensweise:
Stationäre Stelle ermitteln
=> Austauschverhältnis ermitteln
=> Berechnen der Extrema unter Verwendung der Nebenbedingung

Anmerkung: notwendige Bedingung genügt bei konvexen Zielfunktionen bzw. Nebenbedingungen.

Für den Fall n Variablen und k Nebenbedingungen gilt folgende Vorgehensweise:

$$L(x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_k) = f(x_1, \ldots, x_n) + \sum_{i=1}^{k} \lambda_i \cdot g_i(x_1, \ldots, x_n)$$

$f(x_1, \ldots, x_n)$ *ist die Zielfunktion* $\quad g_i(x_1, \ldots, x_n)$ sind die Nebenbedingungen

Daraus entsteht im ersten Durchgang folgende erweiterte Hesse-Matrix:

$$
H_{L(x_1,\dots,x_n,\lambda_1,\dots,\lambda_k)} = \begin{array}{cc}
\begin{array}{ccccccc} \lambda_1 & \lambda_2 & \dots & \lambda_k & x_1 & \dots & x_n \end{array} & \\
\left(\begin{array}{ccccccc}
L_{\lambda_1,\lambda_1} & L_{\lambda_1,\lambda_2} & \dots & L_{\lambda_1,\lambda_k} & L_{\lambda_1,x_1} & \dots & L_{\lambda_1,x_n} \\
L_{\lambda_2,\lambda_1} & L_{\lambda_2,\lambda_2} & \dots & L_{\lambda_2,\lambda_k} & L_{\lambda_2,x_1} & \dots & L_{\lambda_2,x_n} \\
\dots & \dots & \dots & \dots & \dots & \dots & \dots \\
L_{\lambda_k,\lambda_1} & L_{\lambda_k,\lambda_2} & \dots & L_{\lambda_k,\lambda_k} & L_{\lambda_k,x_1} & \dots & L_{\lambda_k,x_n} \\
L_{x_1,\lambda_1} & L_{x_1,\lambda_2} & \dots & L_{x_1,\lambda_k} & L_{x_1,x_1} & \dots & L_{x_1,x_n} \\
\dots & \dots & \dots & \dots & \dots & \dots & \dots \\
L_{x_n,\lambda_1} & L_{x_n,\lambda_2} & \dots & L_{x_n,\lambda_k} & L_{x_n,x_1} & \dots & L_{x_n,x_n}
\end{array}\right) &
\begin{array}{c} \lambda_1 \\ \lambda_2 \\ \dots \\ \lambda_k \\ x_1 \\ \dots \\ x_n \end{array}
\end{array}
$$

Nun können einige Vereinfachungen durchgeführt werden:

(1) Nach dem Satz von Schwarz gilt: $L_{\lambda_a,x_b} = L_{x_b,\lambda_a}$ **=> Matrix ist symmetrisch**

(2) Für die Ableitung nach λ_i **gilt:**

Ableitung der Lagrangefunktion nach den Lagrangemultiplikatoren

$\rightarrow$ erste partielle Ableitung:

$$L_{\lambda_i}(x_1,\dots,x_n,\lambda_1,\dots,\lambda_k) = \frac{\partial L}{\partial \lambda_i}(x_1,\dots,x_n,\lambda_1,\dots,\lambda_k) =$$

$$L_{\lambda_i} = \frac{\partial L}{\partial \lambda_i} f(x_1,\dots,x_n) + \frac{\partial L}{\partial \lambda_i} \lambda_1 \cdot g_1(x_1,\dots,x_n)$$

$$+ \dots + \frac{\partial L}{\partial \lambda_i} \lambda_i \cdot g_i(x_1,\dots,x_n) + \dots + \frac{\partial L}{\partial \lambda_i} \lambda_k \cdot g_k(x_1,\dots,x_n)$$

$$L_{\lambda_i} = 0 + 0 + \dots + g_i(x_1,\dots,x_n) + \dots + 0$$

Nochmaliges Ableiten nach den Lagrangemultiplikatoren:

$$L_{\lambda_i\lambda_i} = \frac{\partial^2 L}{\partial \lambda_i \partial \lambda_i} g_i(x_1,\dots,x_n) = 0$$

Zudem gilt:

$$L_{\lambda_i x_i} = \frac{\partial^2 L}{\partial \lambda_i \partial x_i} g_i(x_1,\dots,x_n) = g_{i_{x_i}}(x_1,\dots,x_n)$$

Ergebnis: die durch die partiellen Ableitungen vereinfachte Hessematrix:

$$H_{L(x_1,\ldots,x_n,\lambda_1,\ldots,\lambda_k)} = \left(\begin{array}{ccccccc|c}
\lambda_1 & \lambda_2 & \ldots & \lambda_k & x_1 & \ldots & x_n & \\ \hline
0 & 0 & \ldots & 0 & g_{1,x_1} & \ldots & g_{1,x_n} & \lambda_1 \\
0 & 0 & \ldots & 0 & g_{2,x_1} & \ldots & g_{2,x_n} & \lambda_2 \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
0 & 0 & \ldots & 0 & g_{k,x_1} & \ldots & g_{k,x_n} & \lambda_k \\
g_{1,x_1} & g_{2,x_1} & \ldots & g_{k,x_1} & L_{x_1,x_1} & \ldots & L_{x_1,x_n} & x_1 \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
g_{1,x_n} & g_{2,x_n} & \ldots & g_{k,x_n} & L_{x_n,x_1} & \ldots & L_{x_n,x_n} & x_n
\end{array}\right)$$

Notwendiges Kriterium für eine Extremwertstelle unter Nebenbedingungen:

Gegeben sei die Zielfunktion $f\left(\vec{x}\right) = f(x_1,\ldots,x_n)$ **mit n unabhängigen Variablen** $\vec{x} = (x_1,\ldots,x_n)$ **und** $g_i(x_1,\ldots,x_n)$ mit $i \in \{1;2;\ldots;k\}$ k Nebenbedingungen

$$\Rightarrow L(x_1,\ldots,x_n,\lambda_1,\ldots,\lambda_k) = f(x_1,\ldots,x_n) + \sum_{i=1}^{k} \lambda_i \cdot g_i(x_1,\ldots,x_n)$$

die in $\overline{x} = \left(\overline{x_1},\ldots,\overline{x_n}\right)$ **bzw.** $\overline{S_L} = \left(\overline{x},\overline{\lambda}\right) = \left(\overline{x_1},\ldots,\overline{x_n},\overline{\lambda_1},\ldots,\overline{\lambda_k}\right)$ **eine stationäre Stelle besitzt.**

Hinreichendes Kriterium für eine Extremwertstelle unter Nebenbedingungen:

Neben der notwendigen Bedingung muss

$L(x_1,\ldots,x_n,\lambda_1,\ldots,\lambda_k) = f(x_1,\ldots,x_n) + \sum_{i=1}^{k} \lambda_i \cdot g_i(x_1,\ldots,x_n)$ **zweimal stetig differenzierbar sein und folgende Bedingung(en) erfüllen:**

Bei Betrachtung der Determinanten der Hauptminoren bzw. aller Haupunterdeterminanten der erweiterten Hesse-Matrix mit eingesetzten stationären Stellen
$\overline{S_L} = \left(\overline{x},\overline{\lambda}, f\right) = \left(\overline{x_1},\ldots,\overline{x_n},\overline{\lambda_1},\ldots,\overline{\lambda_k}, f\right)$**,**

$$H_{L(x_1,\dots,x_n,\lambda_1,\dots,\lambda_k)} = \begin{array}{cc} \begin{matrix} \lambda_1 & \lambda_2 & \dots & \lambda_k & x_1 & \dots & x_n \end{matrix} & \\ \begin{pmatrix} 0 & 0 & \dots & 0 & g_{1,x_1} & \dots & g_{1,x_n} \\ 0 & 0 & \dots & 0 & g_{2,x_1} & \dots & g_{2,x_n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & g_{k,x_1} & \dots & g_{k,x_n} \\ g_{1,x_1} & g_{2,x_1} & \dots & g_{k,x_1} & L_{x_1,x_1} & \dots & L_{x_1,x_n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ g_{1,x_n} & g_{2,x_n} & \dots & g_{k,x_n} & L_{x_n,x_1} & \dots & L_{x_n,x_n} \end{pmatrix} & \begin{matrix} \lambda_1 \\ \lambda_2 \\ \dots \\ \lambda_k \\ x_1 \\ \dots \\ x_n \end{matrix} \end{array}$$

welche mehr als 2k Zeilen/Spalten haben

MINIMUM $\Leftrightarrow$ **Alle Determinanten besitzen das Vorzeichen** $(-1)^k$

Sollten einige (wenige) Det. den Wert 0 besitzen, werden diese ignoriert.

MAXIMUM $\Leftrightarrow$ **Alle Determinanten besitzen wechselnde Vorzeichen mit**

$$Det\left[H\left(\overline{S_L}\right)\right] \quad Vorzeichen \quad (-1)^n$$

Kein Extremum: ***Determinantenwerte besitzen nicht die vorgegebene VZ-Folge.***

Anlagen/Quellen:

https://mjo.osborne.economics.utoronto.ca/index.php/tutorial/index/1/toc

http://www.slimy.com/~steuard/teaching/tutorials/Lagrange.html

http://web.cs.iastate.edu/~cs577/handouts/lagrange-multiplier.pdf

https://www.uni-marburg.de/de/fb02/professuren/qm/statistik/links

Beispiele:

Fall 1: zwei unabhängige Variablen – eine NB

⇨ n = 2 und k = 1

$$L\left(x_1,x_2,\lambda\right) = f\left(x_1,x_2\right)+\lambda\cdot g\left(x_1,x_2\right) = x_1+x_2+\lambda\cdot\left(1-x_1^{\ 2}-x_2^{\ 2}\right)$$

$$L\left(x_1,x_2,\lambda\right) = f\left(x_1,x_2\right)+\lambda\cdot g\left(x_1,x_2\right) = x_1+x_2+\lambda\cdot\left(1-x_1^{\ 2}-x_2^{\ 2}\right)$$

$$\left.\begin{aligned} L_{x_1}\left(x_1,x_2,\lambda\right) &= 1-2\lambda\cdot x_1=0 \quad => \quad \lambda=\frac{1}{2x_1}\\ L_{x_2}\left(x_1,x_2,\lambda\right) &= 1-2\lambda\cdot x_2=0 \quad => \quad \lambda=\frac{1}{2x_2}\end{aligned}\right\}\frac{1}{2x_1}=\frac{1}{2x_2} \quad => \quad x_1=x_2$$

$$L_\lambda\left(x_1,x_2,\lambda\right) = 1-x_1^{\ 2}-x_1^{\ 2}=0 \quad => \quad 1=2x_1^{\ 2} \quad => \quad x_1=\pm\frac{1}{\sqrt{2}}=x_2$$

$$\lambda=\frac{1}{2\cdot\frac{1}{\sqrt{2}}}=\frac{1}{\sqrt{2}}$$

$$H=\begin{pmatrix} 0 & -2x_1 & -2x_2\\ -2x_1 & -2\lambda & 0\\ -2x_2 & 0 & -2\lambda\end{pmatrix}$$

$$H_1=\begin{pmatrix} 0 & -\sqrt{2} & -\sqrt{2}\\ -\sqrt{2} & -2\cdot\frac{1}{\sqrt{2}} & 0\\ -\sqrt{2} & 0 & -2\cdot\frac{1}{\sqrt{2}}\end{pmatrix}=\begin{pmatrix} 0 & -\sqrt{2} & -\sqrt{2}\\ -\sqrt{2} & -\sqrt{2} & 0\\ -\sqrt{2} & 0 & -\sqrt{2}\end{pmatrix}$$

$$H_2=\begin{pmatrix} 0 & \sqrt{2} & \sqrt{2}\\ \sqrt{2} & \sqrt{2} & 0\\ \sqrt{2} & 0 & \sqrt{2}\end{pmatrix}$$

Auswertung ab 3 Spalten/Zeilen

$$H_1 = \begin{pmatrix} 0 & -\sqrt{2} & -\sqrt{2} \\ -\sqrt{2} & -\sqrt{2} & 0 \\ -\sqrt{2} & 0 & -\sqrt{2} \end{pmatrix}$$

$$\Rightarrow \quad Det \begin{pmatrix} 0 & -\sqrt{2} & -\sqrt{2} \\ -\sqrt{2} & -\sqrt{2} & 0 \\ -\sqrt{2} & 0 & -\sqrt{2} \end{pmatrix} = 0+0+0-\left(-2\cdot\sqrt{2}\right)-\left(-2\cdot\sqrt{2}\right) = 4\cdot\sqrt{2} > \quad \Rightarrow \quad Max$$

$$H_2 = \begin{pmatrix} 0 & \sqrt{2} & \sqrt{2} \\ \sqrt{2} & \sqrt{2} & 0 \\ \sqrt{2} & 0 & \sqrt{2} \end{pmatrix}$$

$$\Rightarrow \quad Det \begin{pmatrix} 0 & \sqrt{2} & \sqrt{2} \\ \sqrt{2} & \sqrt{2} & 0 \\ \sqrt{2} & 0 & \sqrt{2} \end{pmatrix} = 0+0+0-\left(2\cdot\sqrt{2}\right)-\left(2\cdot\sqrt{2}\right) = -4\cdot\sqrt{2} < 0 \quad \Rightarrow \quad Min$$

Fall 2: drei unabhängige Variablen – eine NB

⇨ n = 3 und k = 1

$$L\left(x_1, x_2, x_3, \lambda\right) \quad = \quad f\left(x_1, x_2, x_3\right) + \lambda \cdot g\left(x_1, x_2, x_3\right)$$

$$L\left(x_1, x_2, x_3, \lambda\right) \quad = \quad 2x_1 + 4x_2 + \sqrt{28}x_3 + \lambda \cdot \left(12 - x_1^{\;2} - x_2^{\;2} - x_3^{\;2}\right)$$

Fall 3: drei unabhängige Variablen – zwei NB

⇨ n = 3 und k = 2

$$L\left(x_1, x_2, x_3, \lambda_1, \lambda_2\right) \quad = \quad f\left(x_1, x_2, x_3\right) + \lambda_1 \cdot g_1\left(x_1, x_2, x_3\right) + \lambda_2 \cdot g_2\left(x_1, x_2, x_3\right)$$

$$L\left(x_1, x_2, x_3, \lambda\right) \quad = \quad 2x_1 + 4x_2 + x_3 + \lambda_1 \cdot \left(12 - x_1^{\;2} - x_2^{\;2} - x_3^{\;2}\right) + \lambda_2 \cdot \left(12 - 4x_1 - 2x_2\right)$$