

DAS OPTIMIERUNGSVERFAHREN MIT LAGRANGE-MULTIPLIKATOREN

ROBERT KOSCHIG (WWW.MASSMATICS.DE), 09/2012

Inhaltsverzeichnis

1 Motivation	2
1.1 Wo taucht so etwas auf?	2
1.2 Was ist das Problem?	3
2 Das Lagrange-Verfahren	3
2.1 Der allgemeine Weg	3
2.2 Beispiel	5
2.3 Plus oder Minus Lambda?	8
2.4 Zusammenfassung	11
3 Die Art der Extrema	11
3.1 Grafisches Lösen	12
3.2 Einsetzen/Vergleichen	15
3.3 Die Geränderte Matrix	16
3.3.1 Die Determinante der geränderten Matrix - Allgemein	16
3.3.2 Fall: Zwei Variablen, eine Nebenbedingung	18
3.3.3 Fall: Drei Variablen, eine Nebenbedingung	20
3.3.4 Fall: Drei Variablen, zwei Nebenbedingungen	21
3.4 Hesse-Matrix der Lagrange Funktion	22
4 Literatur und Links	24

1 Motivation

Bei Optimierungsproblemen versucht man eine gegebene Zielfunktion $f(x)$ zu maximieren oder zu minimieren - je nach Aufgabenstellung. Soweit könnte man sich ja eigentlich der herkömmlichen Berechnung von Extremstellen bedienen, die man noch vom Abitur her kennt:

1. Die erste Ableitung bilden
2. Über deren Nullstellen die potenziellen Optimalstellen berechnen und
3. Mit der zweiten Ableitung die Art dieser Extremstelle bestimmen

Gibt es mehrere Maxima (bzw. Minima), dann wählt man für die Lösung des Maximierungs- bzw. Minimierungsproblems den größten bzw. kleinsten Funktionswert der in Frage kommenden Stellen und vergleicht diese ggf. noch mit den Funktionswerten von etwaigen Randpunkten. Im Studium kommt dann meist eine erste Steigerung derartiger Probleme durch die Einführung von Zielfunktionen mit mehreren Variablen wie $f(x, y, z)$. Doch so wild ist das nun auch wieder nicht - der Ablauf bleibt derselbe:

1. Die ersten partiellen Ableitungen bilden,
2. Über deren Nullstellen die potenziellen Optimalstellen bestimmen (hier hat man dann ein Gleichungssystem zu lösen) und
3. Mit der Hesse-Matrix¹ die Art dieser Extrema ermitteln bzw.
4. Per Vergleich der Funktionswerte den größten/kleinsten Wert und so die Optimallösung bestimmen.

Problematisch wird dieses Vorgehen, wenn neben mehreren Variablen zusätzlich noch Restriktionen gegeben sind, z.Bsp. dass nur positive Optimalstellen erlaubt sind oder die Variablen zusätzlich noch eine oder mehrere Gleichungen erfüllen sollen - sogenannte Nebenbedingungen.

1.1 Wo taucht so etwas auf?

Prinzipiell tauchen Optimierungsprobleme mit Nebenbedingungen in allen möglichen Fachgebieten auf. In der Mikroökonomie beschäftigt man sich zum Beispiel meist mit der Nutzenmaximierung unter der Nebenbedingung der Budgetrestriktion. Je nach Wahl der Variablen gibt der Wert des Lagrange Multiplikators dann den Grenznutzen des Einkommens wieder. Doch auch bei Produktions- oder Gewinnfunktionen ist man ja oft daran interessiert Minima bzw. Maxima unter bestimmten Bedingungen zu berechnen. Weitere Anwendungen gehen von der Portfoliotheorie zur optimalen Anlagestrategie bis zur Optimierung im Energieversorgungsbereich.

Ursprünglich kommt die Idee der Lagrange-Multiplikatoren jedoch aus der Physik, wo man verschiedene Formen von Bewegungsgleichungen unter Zuhilfenahme der Lagrange-Multiplikatoren formuliert. Dementsprechend findet man hier auch eine Menge Anwendungen des sogenannten Lagrange-Formalismus. Mehr zu diesen Themen und vor allem Beispiele finden sich unter den angegebenen Quelle in [4](#) unter den Links zu Anwendungsbeispielen.

¹Also der Matrix der zweiten partiellen Ableitungen

1.2 Was ist das Problem?

Wenn wir also ein Optimierungsproblem mit Nebenbedingung gegebene haben, die alle in Form einer Gleichung² vorliegen, so gibt es zwei Lösungswege:

1. Einsetzen der Nebenbedingungen in die Zielfunktion
2. Das Optimierungsverfahren mittels *Lagrange - Multiplikatoren*

Das erste Verfahren bietet sich nur an, wenn man jede Nebenbedingung eindeutig nach einer Variable auflösen kann. Die nach einer Variablen umgestellten Gleichungen ersetzt man dann in der Zielfunktion und reduziert dort so die Anzahl der Variablen. Übrig bleibt ein Problem mit der vereinfachten (weil weniger Variablen) Zielfunktion ohne Nebenbedingung, weil diese ja über das Umstellen und Einsetzen „eingearbeitet“ wurden.

Hier kann man dann wie eingangs erwähnt mit den ersten bzw. zweiten (partiellen) Ableitungen die Extremstellen und ihre Art bestimmen.

Im Allgemeinen ist dieser Weg aber leider nicht immer möglich, sodass man ein Verfahren braucht, die Extremstellen unter Berücksichtigung der Nebenbedingungen zu finden: Hier kommt das Lagrange - Verfahren ins Spiel. Dieses Verfahren benutzt auch die ersten (partiellen) Ableitungen um potenzielle Extremstellen zu bestimmen und ist von der Sicht recht intuitiv. Problematisch wird es aber beim Nachweis von Maximums- bzw. Minimumsstellen, denn hier gibt es keinen allgemein gültigen Weg. Ein paar Möglichkeiten dazu werden aber in Abschnitt 3 gezeigt.

Das Lagrange - Verfahren bietet zusammengefasst einen Lösungsweg für Optimierungsprobleme

- Mit einer differenzierbaren Zielfunktion (mit mehreren Unbekannten),
- Nebenbedingungen, die alle als Gleichung vorliegen und sich insbesondere nicht alle nach einer Variable *eindeutig* umstellen lassen (siehe 2.2)

2 Das Lagrange-Verfahren

Die Idee hinter dem von Joseph-Louis Lagrange entwickelten Verfahren ist recht einfach: Man erweitert die Zielfunktion f , indem man jede Nebenbedingung mit einer zusätzlichen Variable multipliziert (Bezeichnung meist $\lambda_1, \lambda_2, \dots$) und diese auf die Zielfunktion addiert. Das Maximum bzw. Minimum unserer Zielfunktion unter den Nebenbedingungen finden wir nun, indem wir das Maximum bzw. Minimum dieser neuen Funktion bestimmen.

Oft scheiden sich die Geister, ob man die multiplizierten Nebenbedingungen addieren oder von der Zielfunktion subtrahieren soll. Wie am Beispiel 2.3 zu sehen, spielt es aber keine Rolle. Bevor wir zu diesem Beispiel kommen, hier zunächst das allgemeine Vorgehen. Lass dich dabei von den vielen Buchstaben nicht verwirren - in den meisten Problemen hat man nur 2-3 Variablen bei 1-2 Nebenbedingungen.

2.1 Der allgemeine Weg

Nehmen wir an, wir haben ein Optimierungsproblem mit n Variablen x_1, x_2, \dots, x_n , bei welchem wir die Zielfunktion (ZF) $f(x_1, x_2, \dots, x_n)$ unter den k Nebenbedingungen

²Es gibt natürlich auch Optimierungsprobleme mit Ungleichungen als Restriktionen. Einen Überblick gibt es dazu - wie sollte es anders sein - auf [Wikipedia](#)

$$\begin{array}{rcl}
\text{NB}_1 : & g_1(x_1, x_2, \dots, x_n) & = 0 \\
\text{NB}_2 : & g_2(x_1, x_2, \dots, x_n) & = 0 \\
& \vdots & \vdots \\
\text{NB}_k : & g_k(x_1, x_2, \dots, x_n) & = 0
\end{array}$$

maximieren wollen (minimieren ist natürlich genauso möglich). Wichtig ist hier, dass alle Gleichungen nach null umgestellt sind.

Dieses Problem ist mathematisch formuliert das Folgende:

$$\begin{array}{l}
f(x_1, x_2, \dots, x_n) \rightarrow \text{MAX!} \\
g_1(x_1, x_2, \dots, x_n) = 0 \\
g_2(x_1, x_2, \dots, x_n) = 0 \\
\vdots \\
g_k(x_1, x_2, \dots, x_n) = 0
\end{array}$$

Die möglichen Optimalstellen dieses Problems findet man durch diese vier Schritte:

Schritt 1: Sofern noch nicht gegeben, stellst du zuerst jede Nebenbedingung (NB) so um, dass du auf einer Seite der Gleichung eine null hast. Damit ist dann jede Nebenbedingung in der Form von oben:

$$\begin{array}{rcl}
\text{NB}_1 : & g_1(x_1, x_2, \dots, x_n) & = 0 \\
\text{NB}_2 : & g_2(x_1, x_2, \dots, x_n) & = 0 \\
& \vdots & \vdots \\
\text{NB}_k : & g_k(x_1, x_2, \dots, x_n) & = 0
\end{array}$$

Schritt 2: Nun wird jeder Ausdruck auf der rechten Seite mit einer neuen Variable multipliziert (wir nennen diese $\lambda_1, \dots, \lambda_k$) und auf die Zielfunktion (f) addiert. Das ergibt unsere Hilfsfunktion - oft auch *Lagrange - Funktion* genannt:

$$L(x_1, \dots, x_n, \lambda_1, \dots, \lambda_k) = \underbrace{f(x_1, \dots, x_n)}_{\text{ZF}} + \lambda_1 \cdot \underbrace{g_1(x_1, \dots, x_n)}_{\text{NB}_1} + \dots + \lambda_k \cdot \underbrace{g_k(x_1, \dots, x_n)}_{\text{NB}_k} \quad (1)$$

Schritt 3: Jetzt bestimmst du die Extremstellen dieser Funktion, indem du nach jeder Variable partiell ableitest, insbesondere auch nach den verschiedenen Lambdas. In unserem Fall sind das dann $n + k$ Ableitungen.

Schritt 4: Jede dieser Ableitungen setzen wir nun null. Das ergibt ein Gleichungssystem mit $k + n$ Gleichungen:

$$\begin{array}{rcl}
\text{I} & \frac{\partial}{\partial x_1} L(x_1, x_2, \dots, x_n, \lambda_1, \lambda_2, \dots, \lambda_k) & = 0 \\
\text{II} & \frac{\partial}{\partial x_2} L(x_1, x_2, \dots, x_n, \lambda_1, \lambda_2, \dots, \lambda_k) & = 0 \\
& \vdots & \vdots \\
\text{n} & \frac{\partial}{\partial x_n} L(x_1, x_2, \dots, x_n, \lambda_1, \lambda_2, \dots, \lambda_k) & = 0 \\
\text{n+1} & \frac{\partial}{\partial \lambda_1} L(x_1, x_2, \dots, x_n, \lambda_1, \lambda_2, \dots, \lambda_k) & = 0 \\
& \vdots & \vdots \\
\text{n+k} & \frac{\partial}{\partial \lambda_k} L(x_1, x_2, \dots, x_n, \lambda_1, \lambda_2, \dots, \lambda_k) & = 0
\end{array}$$

Dieses Gleichungssystem müssen wir nun „nur noch“ lösen und haben dann die Extremstellen. Diese Stellen werden meist „kritische Punkte“ oder „stationäre Stellen“ genannt. Die zugehörigen Lambda-Werte der Lösung dieses Gleichungssystem werden auch *Lagrange-Multiplikatoren* genannt.

Bemerkungen

- Gerade der letzte Schritt: „Lösen des Gleichungssystems“ bereitet oft am meisten Schwierigkeiten, da es sich meist nicht um ein lineares Gleichungssystem handelt und man daher Hilfsmittel wie den Gauß-Algorithmus nicht anwenden kann. Man sollte versuchen, nach Gleichungen Ausschau zu halten, die man nach einer Variable umstellen kann und so das Einsetzungsverfahren anzuwenden. Ebenso möglich ist das Ausklammern von Termen, um dann per Fallunterscheidung der einzelnen Faktoren Werte für die Variablen zu bekommen.
- Bei den Ableitungen nach den verschiedenen Lambdas muss man nicht wirklich rechnen, denn man erhält immer jeweils die entsprechende Nebenbedingung:

$$\frac{\partial}{\partial \lambda_i} L(x_1, x_2, \dots, x_n, \lambda_1, \dots, \lambda_k) = g_i(x_1, x_2, \dots, x_n)$$

Soweit also der theoretische Ablauf. Kommen wir zu einem einfachen Beispiel:

2.2 Beispiel

Nehmen wir an, wir müssten das folgende Optimierungsproblem lösen:

$$\begin{aligned}
f(x, y) &= x + y \rightarrow \text{MAX!} \\
x^2 + y^2 &= 1
\end{aligned}$$

Das heißt wir suchen den größtmöglichen Wert von $f(x, y)$ (ZF) und die zugehörigen Werte (x, y) , die auch die Nebenbedingung (NB)

$$x^2 + y^2 = 1$$

erfüllen. Zunächst stellen wir fest, dass wir die Nebenbedingung nicht eindeutig umstellen können, da sowohl ein Umstellen nach x :

$$x^2 + y^2 = 1 \Rightarrow x_{1,2} = \pm \sqrt{1 - y^2}$$

als auch ein Umstellen nach y stets zwei verschiedene Lösungen hervorbringen würde:

$$x^2 + y^2 = 1 \Rightarrow y_{1,2} = \pm\sqrt{1-x^2}$$

Daher müssen wir das Verfahren von Lagrange anwenden. Dazu sind folgende vier Schritte nötig:

Schritt 1 Zuerst stellst du die Nebenbedingung so um, dass du auf einer Seite der Gleichung eine null hast. In unserem Beispiel müssen wir die 1 auf die linke Seite holen:

$$g_1(x, y) = x^2 + y^2 - 1 = 0$$

Damit steht das Problem formal richtig da:

$$\begin{aligned} f(x, y) &= x + y \rightarrow \text{MAX!} \\ x^2 + y^2 - 1 &= 0 \end{aligned}$$

Schritt 2: Jetzt multiplizieren wir die Nebenbedingung ($g_1(x, y)$) mit einer neuen Variable (λ) und addieren diesen Ausdruck auf die Zielfunktion. Das ergibt unsere Lagrangefunktion:

$$L(x, y, \lambda) = \underbrace{x + y}_{f(x,y)} + \lambda * \underbrace{(x^2 + y^2 - 1)}_{g_1(x,y)}$$

Schritt 3: Nun bestimmst du die Extremstellen dieser Funktion, indem du nach jeder Variable partiell ableitest. In unserem Fall sind das dann 3 Ableitungen. Einmal nach x ...

$$\begin{aligned} \frac{\partial L(x, y, \lambda)}{\partial x} &= \underbrace{1}_{= \frac{\partial}{\partial x} f(x,y)} + \underbrace{2\lambda x}_{= \frac{\partial}{\partial x} \lambda \cdot g_1(x,y)} \\ &= 1 + 2\lambda x \end{aligned}$$

...einmal nach y :

$$\begin{aligned} \frac{\partial L(x, y, \lambda)}{\partial y} &= \underbrace{1}_{= \frac{\partial}{\partial y} f(x,y)} + \underbrace{2\lambda y}_{= \frac{\partial}{\partial y} \lambda \cdot g_1(x,y)} \\ &= 1 + 2\lambda y \end{aligned}$$

und schließlich nach λ :

$$\begin{aligned} \frac{\partial L(x, y, \lambda)}{\partial \lambda} &= \underbrace{0}_{= \frac{\partial}{\partial \lambda} f(x,y)} + \underbrace{(x^2 + y^2 - 1)}_{= \frac{\partial}{\partial \lambda} \lambda \cdot g_1(x,y)} \\ &= x^2 + y^2 - 1 \end{aligned}$$

Wie wir sehen ist das (wie bei den Bemerkungen angesprochen) wieder unsere Nebenbedingung $g_1(x, y)$.

Schritt 4: Für die gesuchten Extremstellen setzen wir nun jede dieser Ableitungen gleich null. Das ergibt folgendes Gleichungssystem:

$$\begin{array}{lcl} \text{I} & 1 + 2\lambda x & = 0 \\ \text{II} & 1 + 2\lambda y & = 0 \\ \text{III} & x^2 + y^2 - 1 & = 0 \end{array}$$

Um das zu lösen, stellen wir die erste Gleichung nach x um:

$$\begin{aligned} 1 + 2\lambda x &= 0 \\ 2\lambda x &= -1 \\ x &= -\frac{1}{2\lambda} \end{aligned}$$

Durch λ dürfen wir an dieser Stelle teilen, weil λ nicht null sein kann³.

Genauso gehen wir in Gleichung II vor:

$$\begin{aligned} 1 + 2\lambda y &= 0 \\ 2\lambda y &= -1 \\ y &= -\frac{1}{2\lambda} \end{aligned}$$

Beide umgestellten Gleichungen können wir nun in die dritte Gleichung einsetzen und so nach λ umstellen:

$$\begin{aligned} x^2 + y^2 - 1 &= 0 \\ \left(-\frac{1}{2\lambda}\right)^2 + \left(-\frac{1}{2\lambda}\right)^2 - 1 &= 0 \\ \frac{1}{4\lambda^2} + \frac{1}{4\lambda^2} &= 1 \\ 2 &= 4\lambda^2 \\ \lambda^2 &= \frac{1}{2} \\ \lambda_1 &= -\frac{1}{\sqrt{2}} \\ \lambda_2 &= \frac{1}{\sqrt{2}} \end{aligned}$$

Jeden dieser beiden Lambda-Werte setzen wir nun in die Gleichung für x bzw. y ein und bekommen so zwei mögliche Optimalstellen:

$$\begin{array}{lcl} x_1 & = -\frac{1}{2\lambda_1} & y_1 = -\frac{1}{2\lambda_1} \\ & = -\frac{1}{2 \cdot \left(-\frac{1}{\sqrt{2}}\right)} & = -\frac{1}{2 \cdot \left(-\frac{1}{\sqrt{2}}\right)} \\ & = \frac{\cancel{\sqrt{2}}}{\cancel{2}} & = \frac{\cancel{\sqrt{2}}}{\cancel{2}} \\ & = \frac{1}{\sqrt{2}} & = \frac{1}{\sqrt{2}} \end{array}$$

³Dann stünde in der Gleichung $1 + 0 = 0$ und das wäre nicht lösbar

Auch bei der zweiten Lösung gilt $x_2 = y_2$:

$$\begin{aligned} x_2 &= -\frac{1}{2\lambda_2} & y_2 &= -\frac{1}{2\lambda_2} \\ &= -\frac{1}{2 \cdot \frac{1}{\sqrt{2}}} & &= -\frac{1}{2 \cdot \frac{1}{\sqrt{2}}} \\ &= -\frac{\sqrt{2}}{2} & &= -\frac{\sqrt{2}}{2} \\ &= -\frac{1}{\sqrt{2}} & &= -\frac{1}{\sqrt{2}} \end{aligned}$$

Am Ende stehen also die beiden potenziellen Optimalstellen

$$E_1 \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right) \text{ und } E_2 \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right)$$

Wie man allgemein bestimmt, welche dieser Stellen nun die Optimallösung darstellt ist keine leichte Frage - „nicht trivial“ wie man so schön sagt - und wird im Abschnitt 3 behandelt. An dieser Stelle könnte man einfach die Werte der *Zielfunktion*

$$f(x, y) = x + y$$

für die beiden Stellen vergleichen (wohlgemerkt ohne λ_1 bzw. λ_2) und schauen, welcher der größere ist. Bei unserem Beispiel ist relativ klar, dass das E_1 ist:

$$\begin{aligned} f \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right) &= \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} \\ &= \frac{2}{\sqrt{2}} \\ &= \sqrt{2} \\ f \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right) &= -\frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}} \\ &= -\frac{2}{\sqrt{2}} \\ &= -\sqrt{2} \end{aligned}$$

Im letzten Schritt haben wir hier benutzt, dass

$$-\frac{2}{\sqrt{2}} = -\frac{\sqrt{2}\sqrt{2}}{\sqrt{2}}$$

An der Stelle $\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right)$ wird also unsere Zielfunktion unter der Nebenbedingung $x^2 + y^2 = 1$ maximiert. Minimiert wird die Zielfunktion bei $\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right)$.

2.3 Plus oder Minus Lambda?

Um zu zeigen, dass es keine Rolle spielt, ob man die Nebenbedingungen addiert oder subtrahiert, wollen wir das Beispiel von eben mal mit der Alternative durchrechnen, dass wir die Nebenbedingung abziehen. Im zweiten Schritt hätten wir also folgende Lagrange - Funktion:

$$\tilde{L}(x, y, \lambda) = \underbrace{x + y}_{f(x,y)} \boxed{-} \lambda * \underbrace{(x^2 + y^2 - 1)}_{g_1(x,y)}$$

Schritt 3: Wenn wir diese nun ableiten, bekommen wir für die partielle Ableitung nach x :

$$\begin{aligned} \frac{\partial}{\partial x} \tilde{L}(x, y, \lambda) &= \underbrace{1}_{=\frac{\partial}{\partial x} f(x,y)} - \underbrace{2\lambda x}_{=\frac{\partial}{\partial x} \lambda \cdot g_1(x,y)} \\ &= 1 - 2\lambda x \end{aligned}$$

Das Minus - Zeichen bleibt genauso bei der Ableitung nach y erhalten:

$$\begin{aligned} \frac{\partial}{\partial y} \tilde{L}(x, y, \lambda) &= \underbrace{1}_{=\frac{\partial}{\partial y} f(x,y)} - \underbrace{2\lambda y}_{=\frac{\partial}{\partial y} \lambda \cdot g_1(x,y)} \\ &= 1 - 2\lambda y \end{aligned}$$

Auch bei der Ableitung nach λ bekommt man nun die Nebenbedingung mit zusätzlichem Minus-Zeichen:

$$\begin{aligned} \frac{\partial}{\partial \lambda} \tilde{L}(x, y, \lambda) &= \underbrace{0}_{=\frac{\partial}{\partial \lambda} f(x,y)} - \underbrace{(x^2 + y^2 - 1)}_{=\frac{\partial}{\partial \lambda} \lambda \cdot g_1(x,y)} \\ &= -x^2 - y^2 + 1 \end{aligned}$$

Schritt 4: Das Gleichungssystem, was wir somit erhalten, lautet:

$$\begin{array}{lcl} \text{I} & 1 - 2\lambda x & = 0 \\ \text{II} & 1 - 2\lambda y & = 0 \\ \text{III} & -x^2 - y^2 + 1 & = 0 \end{array}$$

Wie eben, lösen wir die ersten beiden Gleichungen nach x bzw. y auf und erhalten:

$$x = \frac{1}{2\lambda} \text{ bzw. } y = \frac{1}{2\lambda}$$

Das in die dritte Gleichung eingesetzt ergibt:

$$\begin{aligned}
-x^2 - y^2 + 1 &= 0 \\
-\left(\frac{1}{2\lambda^2}\right)^2 - \left(\frac{1}{2\lambda}\right)^2 + 1 &= 0 \\
-\frac{1}{4\lambda^2} - \frac{1}{4\lambda^2} &= -1 \\
-2 &= -4\lambda^2 \\
\lambda^2 &= \frac{1}{2} \\
\lambda_1 &= -\frac{1}{\sqrt{2}} \\
\lambda_2 &= \frac{1}{\sqrt{2}}
\end{aligned}$$

Wenn wir das mit den Lösungen von oben (2.2) vergleichen, sehen wir, dass das schon mal dasselbe ist. Wir setzen λ_1 in die Gleichung für x bzw. y ein und bekommen:

$$\begin{aligned}
x_1 &= \frac{1}{2\lambda_1} & y_1 &= \frac{1}{2\lambda_1} \\
&= \frac{1}{2 \cdot \left(-\frac{1}{\sqrt{2}}\right)} & &= \frac{1}{2 \cdot \left(-\frac{1}{\sqrt{2}}\right)} \\
&= -\frac{\cancel{\sqrt{2}}}{\cancel{2}} & &= -\frac{\cancel{\sqrt{2}}}{\cancel{2}} \\
&= -\frac{1}{\sqrt{2}} & &= -\frac{1}{\sqrt{2}}
\end{aligned}$$

Analog bekommen wir auch die zweite Lösung:

$$\begin{aligned}
x_2 &= \frac{1}{2\lambda_2} & y_2 &= \frac{1}{2\lambda_2} \\
&= \frac{1}{2 \cdot \left(\frac{1}{\sqrt{2}}\right)} & &= \frac{1}{2 \cdot \left(\frac{1}{\sqrt{2}}\right)} \\
&= \frac{\cancel{\sqrt{2}}}{\cancel{2}} & &= \frac{\cancel{\sqrt{2}}}{\cancel{2}} \\
&= \frac{1}{\sqrt{2}} & &= \frac{1}{\sqrt{2}}
\end{aligned}$$

Hier haben wir also die beiden Optimumsstellen:

$$E'_1 \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right) \text{ und } E'_2 \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right)$$

Wenn wir das mit den Lösungen der Variante vergleichen, wo wir die Nebenbedingung *addiert* haben, sehen wir zunächst dass diese Stellen nicht dasselbe sind:

$$E_1 \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right) \text{ und } E_2 \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right)$$

Die Lambda - Werte haben unterschiedliche Vorzeichen (die Nummerierung von E_1 bzw. E_2 spielt ja keine Rolle). Dieser Unterschied ist aber im Hinblick auf unsere Problemlösung

Erstens: Nicht bedeutsam - wir interessieren uns ja nur für die x bzw. y Werte. Das λ ist in gewisser Weise nur eine Hilfsvariable, um die Optimallösung zu finden. Denn auch mit E'_1 bzw. E'_2 bekommen wir das Maximum unserer Zielfunktion bei $(x_e, y_e) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$.

Zweitens: Sinnvoll, da wir ja auch eine unterschiedliche Lagrangefunktion haben. Wenn wir dort jeweils die optimalen Stellen einsetzen bekommen wir jeweils denselben Wert. Die erste Variante unserer Lagrangefunktion (Nebenbedingung addiert) ergibt mit E_1 :

$$\begin{aligned} L(E_1) &= L\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) = \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} + \left(-\frac{1}{\sqrt{2}}\right) \cdot \left(\left(\frac{1}{\sqrt{2}}\right)^2 + \left(\frac{1}{\sqrt{2}}\right)^2 - 1\right) \\ &= \frac{2}{\sqrt{2}} - \left(\frac{1}{2} + \frac{1}{2} - 1\right) \\ &= \sqrt{2} \end{aligned}$$

Dasselbe bekommen wir bei der zweiten Variante (Nebenbedingung subtrahiert) mit E'_2 :

$$\begin{aligned} \tilde{L}(E'_2) &= \tilde{L}\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}} \cdot \left(\left(\frac{1}{\sqrt{2}}\right)^2 + \left(\frac{1}{\sqrt{2}}\right)^2 - 1\right) \\ &= \frac{2}{\sqrt{2}} - \left(\frac{1}{2} + \frac{1}{2} - 1\right) \\ &= \sqrt{2} \end{aligned}$$

Analog ergeben auch $L(E_2)$ und $\tilde{L}(E'_1)$ genau dieselben Werte.

Die unterschiedlichen Vorzeichen beim Addieren bzw. Subtrahieren der Nebenbedingungen haben also keinerlei Einfluss auf das Ergebnis - hier darf dann wohl dein Dozent entscheiden, was „richtig“ ist.

2.4 Zusammenfassung

Zusammengefasst gibt es also vier Punkte, die du dir zum Lösen von Problemen mit Lagrange - Multiplikatoren merken musst:

1. Du musst die Gleichungen der Nebenbedingungen alle nach null umstellen
2. Danach wird jede Nebenbedingung mit einer neuen Variable multipliziert und auf die Zielfunktion addiert (man kann auch subtrahieren - das spielt keine Rolle)
3. Diese neue Hilfsfunktion (Lagrange - Funktion) wird nach allen Variablen (also auch den Neuen aus dem 2. Schritt) abgeleitet
4. Jede Ableitung muss null sein - aus diesen Bedingungen ergibt sich ein Gleichungssystem. Die Lösung dieses Systems sind die gesuchten Optimalstellen

3 Die Art der Extrema

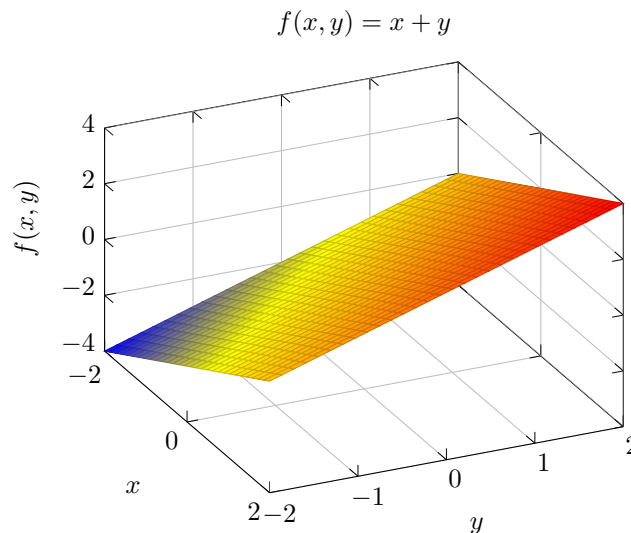
Dieses Vorgehen lässt die Frage offen, ob die gefundene Stelle wirklich das gesuchte Maximum (bzw. Minimum) ist. Spätestens bei mehreren Lösungen braucht man hier ja eine Entscheidungshilfe. Die schlechte

Nachricht ist, dass es im Gegensatz zur normalen Extremwertberechnung, wo man die zweite Ableitung bzw. Hesse - Matrix heranziehen kann, keinen allgemeingültigen Weg gibt. Vielmehr gibt es eine Reihe an Optionen, die wir nun noch vorstellen wollen:

1. Grafisch Lösen/Veranschaulichen (3.1)
2. Einsetzen und Vergleichen (3.2)
3. Geränderte Matrix (3.3)
4. Hesse-Matrix der Lagrange Funktion (3.4)

3.1 Grafisches Lösen

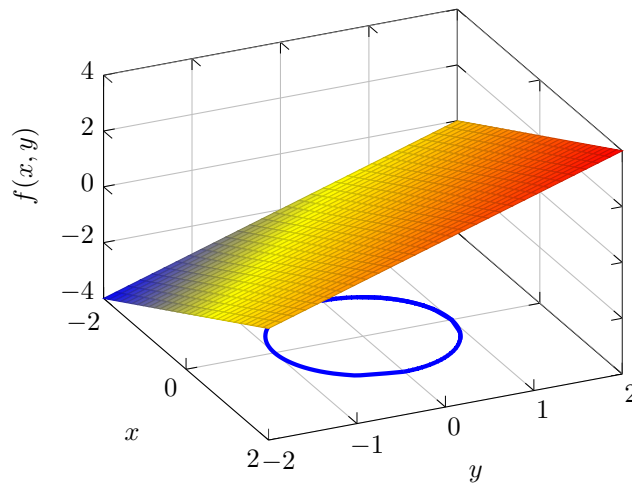
Unser Problem aus dem Beispiel (2.2) lässt sich auch grafisch veranschaulichen. Erste Herausforderung ist dabei die Zeichnung der Zielfunktion $f(x, y) = x + y$:



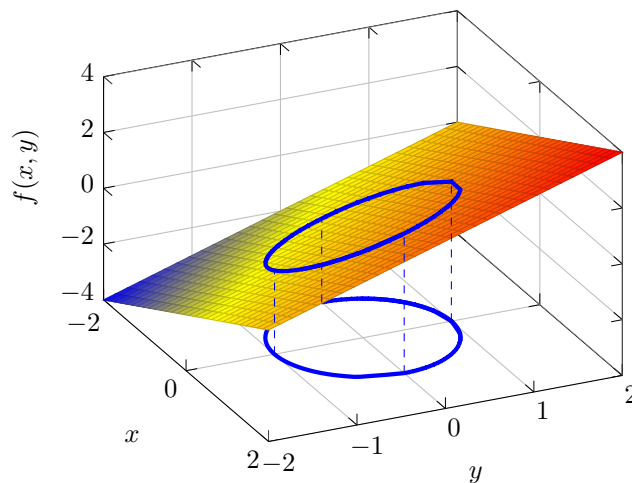
Zu der kommt die Nebenbedingung, dass

$$x^2 + y^2 = 1$$

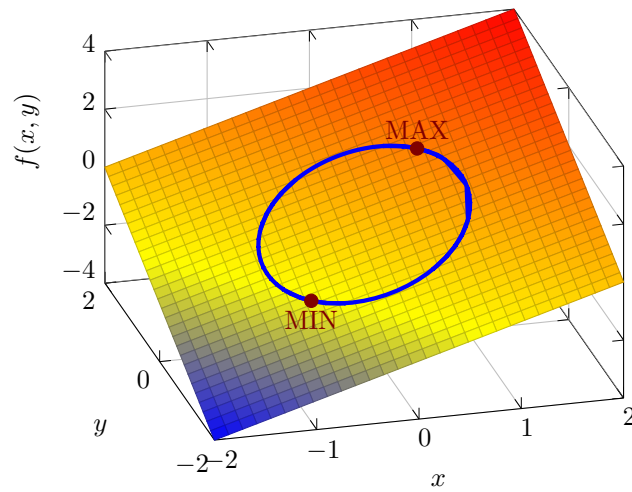
sein soll. Nach etwas Recherche im Tafelwerk kommt man hier darauf, dass diese Gleichung einen Kreis um den Mittelpunkt $M(0,0)$ mit Radius $r = 1$ beschreibt - den sogenannten Einheitskreis. Fügen wir den mal in unserer Zeichnung hinzu:



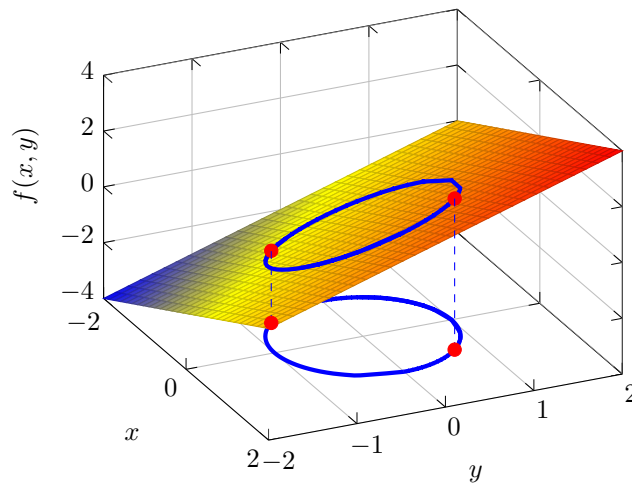
Nun bedeutet die Aufgabe in Worten: „Finde den größten Wert von $f(x, y)$, für den x und y Koordinate auf dem Einheitskreis liegen“ Das heißt, wir zeichnen im nächsten Schritt die Punkte auf unserer Funktionsfläche ein, wo die x und y Koordinate auf dem blauen Kreis liegen. Anders ausgedrückt: Wir projizieren den Einheitskreis auf unsere Funktionsfläche:



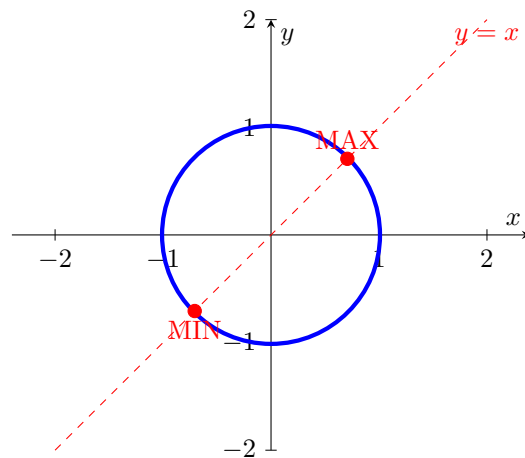
Wenn wir die möglichen Punkte auf unserer Funktionsfläche betrachten, sehen wir schnell, wo die Funktion am höchsten (MAX) bzw. niedrigsten (MIN) ist:



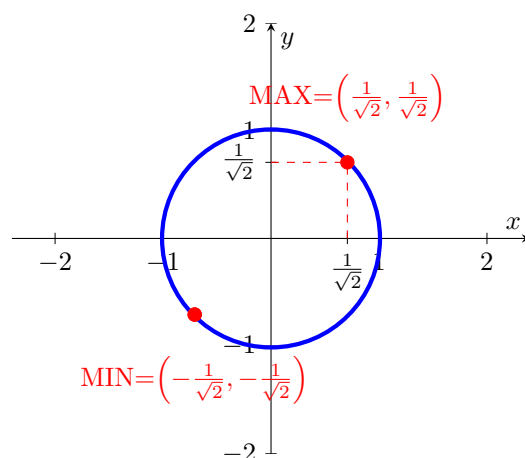
Um die zugehörigen Koordinaten zu bestimmen, kann man nun das Koordinatensystem wieder zurückdrehen...



...und sieht, dass die Punkte für das Maximum bzw. Minimum auf der Geraden liegen, wo $x = y$ ist:



Natürlich haben wir das oben alles schon berechnet: die Punkte auf dem Einheitskreis, wo $x = y$ gilt, sind genau die Lösungen E_1 bzw. E_2 , die wir oben im Beispiel (2.2) erhalten haben:



Damit wäre ein weiteres mal gezeigt, dass das Maximierungsproblem für $(x, y) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ gelöst wird. So schön wie diese Grafiken auch sind, sie dienen wenn dann mehr als Argumentationshilfe, denn:

- Bei mehr als 2 Variablen dürfte es schwer werden auch nur eine Vorstellung des Funktionsbildes zu bekommen
- Mehrere Nebenbedingungen erhöhen den Schwierigkeitsgrad noch zusätzlich
- Oft kann man Zielfunktion und Nebenbedingung auch nicht mehr sinnvoll/aussagekräftig skizzieren, gerade bei fehlenden Hilfsmitteln
- In Klausuren bzw. Prüfungen fehlt zum einen Zeit für solche Kunstwerke - zum anderen gibt es selten verwertbare Anerkennung des Dozenten bzw. Professors für bunte Bilder

Damit also zu den rechnerischen Varianten.

3.2 Einsetzen/Vergleichen

Wie wir eben im Beispiel bereits gesehen haben, kann man natürlich die verschiedenen Optimalstellen einfach in die Zielfunktion einsetzen und dann schauen, welcher Wert am besten (größten bzw. kleinsten) ist. So einfach, wie dieses Vorgehen wirkt, so problematisch ist es auch:

- Was macht man, wenn man nur *eine* Optimalstelle hat - wieso sollte das automatisch das Maximum (bzw. Minimum) sein? Nur weil es so in der Aufgabenstellung steht dürfte nicht als Begründung ausreichen
- Bei vielen Werten muss man dementsprechend viel Rechnen um unter mehreren potenziellen Optimalstellen die größte (bzw. kleinste) herauszufinden

Und natürlich stellt sich die Frage, ob einsetzen und vergleichen von Werten alles ist, was Mathematiker zur Lösung dieses Problems im Angebot haben :)

3.3 Die Geränderte Matrix

Ein Weg, doch rechnerisch die Art des jeweiligen Extremas zu bestimmen ist die geränderte bzw. berändete Matrix, welche alle zweiten partiellen Ableitungen der Lagrange-Funktion beinhaltet. Berandet bzw. gerändert heißt sie deshalb, weil die ersten Zeilen bzw. Spalten die Ableitungen der Lambdas beinhalten und damit im Endeffekt die Ableitungen der Nebenbedingungen darstellen. Wir stellen erstmal das allgemeine Kriterium vor, um dann im Anschluss die Spezialfälle vorzustellen, die einem so im täglichen Umgang mit Lagrange-Problemen begegnen können.

3.3.1 Die Determinante der geränderten Matrix - Allgemein

Wie beim allgemeinen Vorgehen (2.1) gehen wir davon aus, dass wir n Variablen und k Nebenbedingungen haben, also letztlich folgendes Problem betrachten:

$$\begin{aligned} f(x_1, \dots, x_n) &\rightarrow \text{MAX!} \\ g_1(x_1, \dots, x_n) &= 0 \\ &\vdots \\ g_k(x_1, \dots, x_n) &= 0 \end{aligned}$$

Wir schreiben die Lagrangefunktion nun so auf, dass zuerst die Lambdas kommen. Das kommt daher, dass wir die Hesse-Matrix dieser Funktion auch so notieren wollen, dass die Ableitungen nach Lambda zuerst kommen:

$$L(\lambda_1, \dots, \lambda_k, x_1, \dots, x_n) = f(x_1, \dots, x_n) + \lambda_1 \cdot g_1(x_1, \dots, x_n) + \dots + \lambda_k \cdot g_k(x_1, \dots, x_n)$$

Die Hesse-Matrix dieser Funktion sieht allgemein erstmal so aus:

$$H_L(\lambda_1, \dots, \lambda_k, x_1, \dots, x_n) = \begin{pmatrix} L_{\lambda_1, \lambda_1} & L_{\lambda_1, \lambda_2} & \dots & L_{\lambda_1, \lambda_k} & L_{\lambda_1, x_1} & \dots & L_{\lambda_1, x_n} \\ L_{\lambda_2, \lambda_1} & L_{\lambda_2, \lambda_2} & \dots & L_{\lambda_2, \lambda_k} & L_{\lambda_2, x_1} & \dots & L_{\lambda_2, x_n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ L_{\lambda_k, \lambda_1} & L_{\lambda_k, \lambda_2} & \dots & L_{\lambda_k, \lambda_k} & L_{\lambda_k, x_1} & \dots & L_{\lambda_k, x_n} \\ L_{x_1, \lambda_1} & L_{x_1, \lambda_2} & \dots & L_{x_1, \lambda_k} & L_{x_1 x_1} & \dots & L_{x_1 x_n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ L_{x_n, \lambda_1} & L_{x_n, \lambda_2} & \dots & L_{x_n, \lambda_k} & L_{x_n, x_1} & \dots & L_{x_n x_n} \end{pmatrix}$$

Nun kann man zwei Dinge vereinfachen. Zum einen kann man in der Regel den Satz von Schwarz anwenden, d.h. unsere Matrix ist symmetrisch, denn es gilt für alle gemischten Ableitungen, wie L_{x_1, λ_1} :

$$L_{x_1, \lambda_1} = L_{\lambda_1, x_1}$$

Zum anderen ergibt die Ableitung nach λ_i , wie oben angesprochen, die i -te Nebenbedingung:

$$\begin{aligned}
L_{\lambda_i} &= \frac{\partial}{\partial \lambda_i} L(\lambda_1, \dots, \lambda_k, x_1, \dots, x_n) \\
&= \frac{\partial}{\partial \lambda_i} (f(x_1, \dots, x_n) + \lambda_1 \cdot g_1(x_1, \dots, x_n) + \dots + \lambda_i \cdot g_i(x_1, \dots, x_n) + \dots + \lambda_k \cdot g_k(x_1, \dots, x_n)) \\
&= \underbrace{\frac{\partial}{\partial \lambda_i} f(x_1, \dots, x_n)}_{=0} + \underbrace{\frac{\partial}{\partial \lambda_i} \lambda_1 \cdot g_1(x_1, \dots, x_n)}_{=0} + \dots + \underbrace{\frac{\partial}{\partial \lambda_i} \lambda_i \cdot g_i(x_1, \dots, x_n)}_{=g_i(x_1, \dots, x_n)} + \dots + \underbrace{\frac{\partial}{\partial \lambda_i} \lambda_k \cdot g_k(x_1, \dots, x_n)}_{=0} \\
&= g_i(x_1, \dots, x_n)
\end{aligned}$$

Wenn man diese Nebenbedingung nun nochmal nach irgendeinem Lambda ableitet wird sie null. Die Ableitung nach einer Variable wie x_1 schreiben wir so:

$$\begin{aligned}
L_{\lambda_i, x_1} &= \frac{\partial}{\partial x_1} L_{\lambda_i} \\
&= \frac{\partial}{\partial x_1} g_i(x_1, \dots, x_n) \\
&= g_{i, x_1}
\end{aligned}$$

Damit lässt sich die Hesse-Matrix noch etwas angenehmer notieren:

$$H_L(\lambda_1, \dots, \lambda_k, x_1, \dots, x_n) = \begin{pmatrix} 0 & 0 & \dots & 0 & g_{1, x_1} & \dots & g_{1, x_n} \\ 0 & 0 & \dots & 0 & g_{2, x_1} & \dots & g_{2, x_n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & g_{k, x_1} & \dots & g_{k, x_n} \\ g_{1, x_1} & g_{2, x_1} & \dots & g_{k, x_1} & L_{x_1 x_1} & \dots & L_{x_1 x_n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ g_{1, x_n} & g_{2, x_n} & \dots & g_{k, x_n} & L_{x_1 x_n} & \dots & L_{x_n x_n} \end{pmatrix} \quad (2)$$

Man nimmt nun die jeweiligen Optimalstellen, die man am Ende vom Algorithmus (2.1) bekommen hat in diese Matrix ein. Nennen wir die (bzw. diese, wenn man mehrere Stellen hat) mal:

$$E(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n, \bar{\lambda}_1, \dots, \bar{\lambda}_k) \quad (3)$$

Um zu entscheiden, ob unsere Zielfunktion an dieser Stelle/diesen Stellen maximiert/minimiert wird, gibt es das folgende, hinreichende Kriterium:

Satz 3.1. *Hinreichendes Kriterium für eine Extremstelle unter Nebenbedingungen Gegeben sei die Zielfunktion f mit den k Nebenbedingungen g_1, \dots, g_k , welche in*

$$\bar{x} := (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)$$

eine kritische Stelle⁴ besitzen. Weiterhin sind f, g_1, \dots, g_k in \bar{x} zweimal stetig differenzierbar. Dann gilt unter Verwendung der Lagrange-Multiplikatoren $\bar{\lambda}$ ⁵ folgendes, hinreichendes Kriterium für die Art des Extremas:

⁴Vgl. 2.1 Schritt 4

⁵Siehe ebenfalls 2.1 Schritt 4

Betrachtet werden alle Hauptunterdeterminanten⁶ der geränderten Matrix (2) mit eingesetzter Extremstelle der Lagrangefunktion $E = (\bar{x}, \bar{\lambda}) = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n, \bar{\lambda}_1, \dots, \bar{\lambda}_k)$:

$$H_L(\bar{E}) = \begin{pmatrix} 0 & 0 & \dots & 0 & g_{1x_1}(\bar{x}) & \dots & g_{1x_n}(\bar{x}) \\ 0 & 0 & \dots & 0 & g_{2x_1}(\bar{x}) & \dots & g_{2x_n}(\bar{x}) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & g_{kx_1}(\bar{x}) & \dots & g_{kx_n}(\bar{x}) \\ g_{1x_1}(\bar{x}) & g_{2x_1}(\bar{x}) & \dots & g_{kx_1}(\bar{x}) & L_{x_1x_1}(\bar{\lambda}, \bar{x}) & \dots & L_{x_1x_n}(\bar{\lambda}, \bar{x}) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ g_{1x_n}(\bar{x}) & g_{2x_n}(\bar{x}) & \dots & g_{kx_n}(\bar{x}) & L_{x_1x_n}(\bar{\lambda}, \bar{x}) & \dots & L_{x_nx_n}(\bar{\lambda}, \bar{x}) \end{pmatrix}$$

welche mehr als $2k$ Zeilen bzw. Spalten haben. Es gilt Folgendes:

Minimum Besitzen alle diese Determinanten das Vorzeichen $(-1)^k$, so handelt es sich um ein Minimum.

Maximum Besitzen die Determinanten wechselnde Vorzeichen, so dass die Determinante von $H_L(\bar{E})$ das Vorzeichen $(-1)^n$ hat, so liegt ein Maximum vor.

Kein Extremum Besitzen die Determinante eine von den eben genannten verschiedene Vorzeichenfolge, so liegt kein Extremum vor. Dabei werden Nullen bei den Determinanten ignoriert, d.h. z.Bsp. wenn alle Determinanten das Vorzeichen $(-1)^k$ haben und nur einige null sind, liegt dennoch ein Minimum vor.

Zugegebenermaßen wirkt dieses Kriterium nicht zuletzt durch die Vielzahl an Buchstaben etwas verwirrend. Wir wollen daher nun ein paar Spezialfälle näher erläutern und auch an unserem Beispiel (2.2) nachvollziehen.

3.3.2 Fall: Zwei Variablen, eine Nebenbedingung

Wenn man, wie im Beispiel oben (2.2), zwei Variablen und eine Nebenbedingung hat, vereinfacht sich das Kriterium mit $n = 2$ und $k = 1$ recht stark, da wir ja nur die Unterdeterminanten mit mehr als $2 \cdot k = 2$ Zeilen betrachten.

⁶Auch Hauptminor genannt, siehe [Wikipedia](#)

Da unsere Hesse-Matrix der Lagrangefunktion (2) nur drei Zeilen hat...

$$H_L(x, y, \lambda) = \begin{pmatrix} 0 & g_x & g_y \\ g_x & L_{xx} & L_{xy} \\ g_y & L_{xy} & L_{yy} \end{pmatrix}$$

...betrachten wir also nur die Determinante dieser 3×3 Matrix mit den eingesetzten Werten der Extremstelle $\bar{E} = (\bar{x}, \bar{y}, \bar{\lambda})$. Hier gilt:

Minimum Hat $\det H_L(\bar{x}, \bar{y}, \bar{\lambda})$ das Vorzeichen $(-1)^k = (-1)^1 = -1$, so hat f in \bar{E} ein Minimum.

Maximum Hat $\det H_L(\bar{x}, \bar{y}, \bar{\lambda})$ das Vorzeichen $(-1)^n = (-1)^2 = +1$, so hat f in \bar{E} ein Maximum.

Keine Aussage Gilt $\det H_L(\bar{x}, \bar{y}, \bar{\lambda}) = 0$, so lässt sich keine Aussage über das Vorliegen eines Extremums treffen.

Wir wollen mit diesem Kriterium, die Art des Extrums der beiden kritischen Stellen

$$E_1 \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right) \text{ und } E_2 \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right)$$

erneut überprüfen. Wir benötigen für die Hesse-Matrix folgende Angaben der Lagrange-Funktion:

$$\begin{aligned} L(x, y, \lambda) &= x + y + \lambda \cdot (x^2 + y^2 - 1) \\ L_x(x, y, \lambda) &= 1 + 2\lambda x \\ L_y(x, y, \lambda) &= 1 + 2\lambda y \\ L_{xx}(x, y, \lambda) &= 2\lambda \\ L_{xy}(x, y, \lambda) &= 0 \\ L_{yy}(x, y, \lambda) &= 2\lambda \end{aligned}$$

Sowie die ersten partiellen Ableitung der Nebenbedingung:

$$\begin{aligned} g(x, y) &= x^2 + y^2 - 1 \\ g_x(x, y) &= 2x \\ g_y(x, y) &= 2y \end{aligned}$$

Das ergibt die geränderte Matrix der Lagrange-Funktion:

$$H_L(x, y, \lambda) = \begin{pmatrix} 0 & 2x & 2y \\ 2x & 2\lambda & 0 \\ 2y & 0 & 2\lambda \end{pmatrix}$$

Wir fangen mit $E_1 \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right)$ an und setzen ein. Für die Berechnung der Determinante hilft die Regel von Sarrus:

$$\begin{aligned}
\det\left(H_L\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)\right) &= \det\begin{pmatrix} 0 & 2 \cdot \frac{1}{\sqrt{2}} & 2 \cdot \frac{1}{\sqrt{2}} \\ 2 \cdot \frac{1}{\sqrt{2}} & 2 \cdot \left(-\frac{1}{\sqrt{2}}\right) & 0 \\ 2 \cdot \frac{1}{\sqrt{2}} & 0 & 2 \cdot \left(-\frac{1}{\sqrt{2}}\right) \end{pmatrix} \\
&= \det\begin{pmatrix} 0 & \sqrt{2} & \sqrt{2} \\ \sqrt{2} & -\sqrt{2} & 0 \\ \sqrt{2} & 0 & -\sqrt{2} \end{pmatrix} \\
&= 0 \cdot (-\sqrt{2}) \cdot (-\sqrt{2}) + \sqrt{2} \cdot 0 \cdot \sqrt{2} + \sqrt{2} \cdot \sqrt{2} \cdot 0 \\
&\quad - \sqrt{2} \cdot \sqrt{2} \cdot (-\sqrt{2}) - 0 \cdot 0 \cdot 0 - \sqrt{2} \cdot (-\sqrt{2}) \cdot \sqrt{2} \\
&= 2\sqrt{2} + 2\sqrt{2} \\
&= 4\sqrt{2}
\end{aligned}$$

Das Vorzeichen dieser Determinante ist positiv, also bestätigt sich, dass wir ein Maximum haben. Analog können wir nachweisen, dass in E_2 ein Minimum vorliegt:

$$\begin{aligned}
\det\left(H_L\left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)\right) &= \det\begin{pmatrix} 0 & 2 \cdot \left(-\frac{1}{\sqrt{2}}\right) & 2 \cdot \left(-\frac{1}{\sqrt{2}}\right) \\ 2 \cdot \left(-\frac{1}{\sqrt{2}}\right) & 2 \cdot \frac{1}{\sqrt{2}} & 0 \\ 2 \cdot \left(-\frac{1}{\sqrt{2}}\right) & 0 & 2 \cdot \frac{1}{\sqrt{2}} \end{pmatrix} \\
&= \det\begin{pmatrix} 0 & -\sqrt{2} & -\sqrt{2} \\ -\sqrt{2} & \sqrt{2} & 0 \\ -\sqrt{2} & 0 & \sqrt{2} \end{pmatrix} \\
&= 0 \cdot \sqrt{2} \cdot \sqrt{2} + (-\sqrt{2}) \cdot 0 \cdot (-\sqrt{2}) + (-\sqrt{2}) \cdot (-\sqrt{2}) \cdot 0 \\
&\quad - (-\sqrt{2}) \cdot (-\sqrt{2}) \cdot \sqrt{2} - 0 \cdot 0 \cdot 0 - (-\sqrt{2}) \cdot \sqrt{2} \cdot (-\sqrt{2}) \\
&= -2\sqrt{2} - 2\sqrt{2} \\
&= -4\sqrt{2}
\end{aligned}$$

Das negative Vorzeichen bestätigt also das Minimum. Abschließend sei noch bemerkt, dass die unter 2.3 erwähnte Gleichgültigkeit, ob man nun $+\lambda$ oder $-\lambda$ verwendet auch auf dieses Kriterium zutrifft. Denn wie gesehen, verändern sich dadurch die Vorzeichen der Lambda-Werte in den kritischen Stellen, jedoch wird dieser Vorzeichen auch wieder ausgeglichen, da auch in den partiellen Ableitungen ein zusätzliches Minus-Zeichen auftaucht.

Gehen wir somit weiter zum nächsten Spezialfall.

3.3.3 Fall: Drei Variablen, eine Nebenbedingung

Wie angesprochen, versagt hier jegliche grafische Veranschaulichung, da wir $n = 3$ Variablen haben. Wir bezeichnen die drei Variablen mal mit x, y und z . Dann lautet die geränderte Matrix:

$$H_L(x, y, z, \lambda) = \begin{pmatrix} 0 & g_x & g_y & g_z \\ g_x & L_{xx} & L_{xy} & L_{xz} \\ g_y & L_{xy} & L_{yy} & L_{yz} \\ g_z & L_{xz} & L_{yz} & L_{zz} \end{pmatrix}$$

Hier wollen wir wieder im Punkt $\bar{E} = (\bar{x}, \bar{y}, \bar{z}, \bar{\lambda})$ das Vorliegen eines Maximums/Minimums untersuchen. Wegen $k = 1$ bzw. $2k = 2$, müssen wir diesmal die Determinante der ersten drei Zeilen/Spalten unserer

geränderten Matrix und die Determinante der gesamten Matrix im kritischen Punkt \bar{E} (3) bestimmen. Wir bezeichnen die beiden Matrizen mal mit H_3 bzw. H_4 :

$$\begin{pmatrix} & \color{red}{H_3} & \color{blue}{H_4} \\ \color{red}{\boxed{\begin{matrix} 0 & g_x & g_y & g_z \\ g_x & L_{xx} & L_{xy} & L_{xz} \\ g_y & L_{xy} & L_{yy} & L_{yz} \\ g_z & L_{xz} & L_{yz} & L_{zz} \end{matrix}}} & & \end{pmatrix}$$

Das Kriterium der Hauptunterdeterminante der geränderten Matrix besagt damit Folgendes:

Minimum Sind die Werte von $\det(H_3)$ und $\det(H_4)$ negativ ($(-1)^k = (-1)^1 = -1$), so liegt ein Minimum vor.

Maximum Hat $\det(H_3)$ das Vorzeichen $+$ und $\det(H_4)$ ein Minus als Vorzeichen ($(-1)^3 = -1$), so liegt ein Maximum vor.

Kein Extremum Ist $\det(H_4)$ positiv, so liegt kein Extremum vor, da dieser Wert sowohl für ein Minimum als auch für ein Maximum negativ sein muss.

Keine Aussage Ist $\det(H_4) = 0$, so lässt sich mit dem Kriterium keine Aussage treffen⁷.

3.3.4 Fall: Drei Variablen, zwei Nebenbedingungen

Weil wir nun zwei Nebenbedingungen haben, betrachten wir nur die Hauptunterdeterminanten mit mehr als $2k = 2 \cdot 2 = 4$ Zeilen. Da die geränderte Matrix hier aber nur $n + k = 3 + 2 = 5$ Zeilen hat ist das die Determinante der gesamten geränderten Matrix. Mit x, y, z für die Variablen und λ_1 und λ_2 für die Faktoren der Nebenbedingungen g_1 bzw. g_2 sieht diese so aus:

$$H_L(x, y, z, \lambda_1, \lambda_2) = \begin{pmatrix} 0 & 0 & g_{1x} & g_{1y} & g_{1z} \\ 0 & 0 & g_{2x} & g_{2y} & g_{2z} \\ g_{1x} & g_{2x} & L_{xx} & L_{xy} & L_{xz} \\ g_{1y} & g_{2y} & L_{xy} & L_{yy} & L_{yz} \\ g_{1z} & g_{2z} & L_{xz} & L_{yz} & L_{zz} \end{pmatrix}$$

Wenn wir hier unserem Kriterium von oben folgen (3.3), so bekommen wir für die kritische Stelle $\bar{E} = (\bar{x}, \bar{y}, \bar{z}, \bar{\lambda}_1, \bar{\lambda}_2)$:

Minimum Ist $\det(H_L(\bar{x}, \bar{y}, \bar{z}, \bar{\lambda}_1, \bar{\lambda}_2))$ positiv ($(-1)^k = (-1)^2 = +1$), so liegt in \bar{E} ein Minimum vor.

Maximum Ist $\det(H_L(\bar{x}, \bar{y}, \bar{z}, \bar{\lambda}_1, \bar{\lambda}_2))$ hingegen negativ ($(-1)^n = (-1)^5 = -1$), so liegt in \bar{E} ein Maximum vor.

Keine Aussage Im Fall $\det(H_L(\bar{x}, \bar{y}, \bar{z}, \bar{\lambda}_1, \bar{\lambda}_2)) = 0$ liefert das Kriterium keine Aussage.

⁷Denn man kann ja nicht entscheiden, welcher der drei beschriebenen Fälle vorliegt.

3.4 Hesse-Matrix der Lagrange Funktion

Von Problemen ohne Nebenbedingung ist man gewohnt, die Art der Extrema über die Definitheit der Hesse-Matrix zu bestimmen. Die Frage ist ja, ob sich dieser Ansatz nicht in irgendeiner Weise übertragen kann. Dazu überlegt man, dass wir beim vorliegenden Optimierungsproblem unsere Variablen x_1, \dots, x_n so wählen wollen, dass sie

1. Die Zielfunktion maximieren/minimieren
2. Die Nebenbedingung erfüllen

Bei der berechneten kritischen Stelle (3) $E(\bar{x}_1, \dots, \bar{x}_n, \bar{\lambda}_1, \dots, \bar{\lambda}_k)$ ist der zweite Punkt auf jeden Fall erfüllt. Um zu prüfen, ob nun zumindest lokal ein Maximum bzw. Minimum vorliegt, betrachtet man das Verhalten der Lagrangefunktion (1) für die Variablen um diese optimale Stelle und lässt die Lagrange Multiplikatoren $(\bar{\lambda}_1, \dots, \bar{\lambda}_k)$ fix. Man bildet also die „Lagrange-Funktion mit eingesetzten Lambdas“ (nennen wir die mal \tilde{L}) und bestimmt die Definitheit deren Hesse-Matrix. Da man meist mehrere Stellen hat, bietet sich an zunächst die Hesse-Matrix der Lagrange-Funktion zu bilden und dann die kritische Stelle einzusetzen. Man beschränkt sich dabei aber auf die partiellen Ableitungen der Variablen:

$$H_{\tilde{L}}(x_1, \dots, x_n) = \begin{pmatrix} L_{x_1 x_1} & \dots & L_{x_1 x_n} \\ \vdots & \ddots & \vdots \\ L_{x_n x_1} & \dots & L_{x_n x_n} \end{pmatrix}$$

Gehen wir auch das mal an unserem Beispiel (2.2) durch. Die Lagrangefunktion war dort:

$$L(x, y, \lambda) = x + y + \lambda(x^2 + y^2 - 1)$$

Mit den partiellen Ableitungen nach x bzw. y können wir dann auch die gesuchte Hesse-Matrix $H_{\tilde{L}}(x, y)$ aufstellen:

$$\begin{aligned} \frac{\partial}{\partial x} L(x, y, \lambda) &= 1 + 2\lambda x \\ \frac{\partial}{\partial y} L(x, y, \lambda) &= 1 + 2\lambda y \\ \Rightarrow H_{\tilde{L}}(x, y) &= \begin{pmatrix} 2\lambda & 0 \\ 0 & 2\lambda \end{pmatrix} \end{aligned}$$

Nun setzen wir die kritischen Punkte

$$E_1 \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right) \text{ und } E_2 \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right)$$

ein und erhalten für E_1 :

$$H_{\tilde{L}} \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right) = \begin{pmatrix} 2 * \left(-\frac{1}{\sqrt{2}} \right) & 0 \\ 0 & 2 * \left(-\frac{1}{\sqrt{2}} \right) \end{pmatrix}$$

Wie man sieht hängt diese Matrix gar nicht von den Werten der Variablen \bar{x} bzw. \bar{y} ab, sondern nur vom Wert des Lagrange-Multiplikators $\lambda = -\frac{1}{\sqrt{2}}$ der kritischen Stelle. Das kann durchaus vorkommen. Da die Matrix Diagonalgestalt hat, stehen die Eigenwerte auf der Diagonale - hier sogar nur ein Eigenwert:

$$-\frac{2}{\sqrt{2}} = -\sqrt{2}$$

Da dieser Wert negativ ist, ist die Matrix negativ definit und wir haben ein lokales Maximum vorliegen. Analog bestätigen wir, dass in E_2 ein Minimum ist:

$$H_{\bar{L}}\left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) = \begin{pmatrix} 2 * \frac{1}{\sqrt{2}} & 0 \\ 0 & 2 * \frac{1}{\sqrt{2}} \end{pmatrix}$$

Hier sind alle Eigenwerte positiv, denn wir haben ja nur einen:

$$\frac{2}{\sqrt{2}} = \sqrt{2}$$

Somit ist die Matrix positiv definit und deshalb liegt ein Minimum vor. Wenn man alle Nebenbedingungen als lineare Funktionen vorliegen hat, dann steckt hinter diesem Vorgehen übrigens nichts weiter als dass man die Hesse Matrix der Zielfunktion an der kritischen Stelle auswertet - dort schließt sich also wieder der Kreis zu Optimierungsproblemen ohne Nebenbedingungen.

Hier nochmal die Kurzzusammenfassung für die Bestimmung der Art der Extrema über die Hesse-Matrix der Lagrange Funktion:

1. Bilde die Hesse - Matrix der Lagrangefunktion (1), wobei hier nur die partiellen Ableitungen der Variablen x_1, \dots, x_n zu beachten sind (alle anderen Zeilen/Spalten lassen wir weg)
2. Setze die kritischen Punkte $E(\bar{x}_1, \dots, \bar{x}_n, \bar{\lambda}_1, \dots, \bar{\lambda}_k)$ jeweils in diese Matrix ein
3. Ist die Matrix positiv definit liegt ein Minimum vor. Ist sie negativ definit, so haben wir ein Maximum
4. Bei indefiniten Matrizen kann man in der Regel keine Aussage treffen - oft liegt dann aber auch keine lokale Extremstelle vor

Am Ende nochmal der Hinweis, dass es von allen vier vorgestellten Verfahren keinen Königsweg gibt. Unserer Meinung nach fährt man rechnerisch meist ganz gut mit dem dritten Verfahren. Bekommt man mehrere Minima bzw. Maxima kann man per Einsetzen und Vergleichen dann noch entscheiden, wo das globale Maximum angenommen wird.

4 Literatur und Links

Wer sich für etwas formale Darstellungen und die genauen Beweise der Gültigkeit der getroffenen Aussagen interessiert, findet hierzu natürlich reichlich Material in der Fachliteratur:

1. S. BOYD UND L. VANDENBERGHE: *Convex optimization*. Cambridge University Press, Cambridge, 2004
2. P. KOSMOL: *Optimierung und Approximation*. De Gruyter, Berlin, 2010
3. H. ROMMELFANGER: *Mathematik für Wirtschaftswissenschaftler I*. BI Wissenschaftsverlag, Mannheim/Leipzig/Wien/Zürich, 1992
4. C.P. SIMON UND L. BLUME: *Mathematics for Economists*. W.W. Norton & Company, New York und London, 1994

Ausführungen zu der Art der Extrema findet man hingegen nicht so häufig. Daher an dieser Stelle auch nur der Verweis auf diese beiden Werke:

1. D.P. BERTSEKAS: *Constrained Optimization and Lagrange Multiplier Methods*. Athena Scientific, Belmont (MA), 1996
2. H. ROMMELFANGER: *Mathematik für Wirtschaftswissenschaftler II*. BI Wissenschaftsverlag, Mannheim/Leipzig/Wien/Zürich, 1992

Neben der Literatur gibt es im Internet unter folgenden Links ein paar Erläuterungen. Falls mehr Quellen erforderlich sind, helfen die Suchbegriffe „Hinreichende Bedingung Extremstelle Lagrange“ bzw. „second order sufficient conditions“:

- [Zusammenfassung der Bedingungen für Optimalstellen von Extremwertproblemen mit Nebenbedingungen der Mathematik der University of California, Los Angeles](#)
- [Übersicht zum Lagrange-Verfahren mit Hinweisen zur Bestimmung der Art der Extrema der Iowa State University](#)
- [Vorstellung des Lagrange Verfahrens mit Bedingung zur Bestimmung der Art der Extrema und Anwendung in der Mikroökonomie, Universität Marburg](#)

Weitere Interpretation und Anwendungsbeispiele gibt es hier:

- [Vorstellung des Lagrange Formalismus und physikalischer Anwendungen von F.Unglaub](#)
- [Kurze Übersicht und Anwendungsbeispiel zur Mikroökonomie](#)
- [Tutorial M.J. Osborne, Universität Toronto](#)
- [Tutorial S. Jensen](#)