R. Herzog, G. Müller
Universität Heidelberg

Übung 5 (Lösung)

Ausgabedatum: 11. November 2024
Abgabedatum: 18. November 2024

Hausaufgabe 5.1 (Zusammenhang von Ecken/Extremalpunkten und zulässigen Basisvektoren) 5 Punkte

Es sei P wie in (6.8) ein Polyeder in Normalform, und es gelte $\operatorname{Rang}(A) = m$. Zeigen Sie die Äquivalenz der folgenden Aussagen:

(a) $x \in \mathbb{R}^n$ ist eine Ecke/ein Extremalpunkt von P.

(b) $x \in \mathbb{R}^n$ ist zulässiger Basisvektor von P.

(c) es existiert ein Kostenvektor $c \in \mathbb{R}^n$, sodass $x \in \mathbb{R}^n$ die einzige Optimallösung des Problems

$$\begin{aligned} \text{Minimiere } & c^\intercal x \quad \text{über } x \in \mathbb{R}^n \\ \text{sodass } & Ax = b \\ \text{und } & x \geq 0 \end{aligned}$$

ist.

Beachte: Die Äquivalenz von Aussagen (a) und (b) ist genau die Aussage von Satz 6.19 aus dem Skript.

Lösung.

(a)„⇒"(b) Sei x eine Ecke von P. Wir bezeichnen die Indexmenge, auf der x positiv ist (die Menge der Indizes zu den *inaktiven* Nebenbedingungen) mit

$$\mathcal{I} = \mathcal{I}(x) \coloneqq \{i \in \{1, \ldots, n\} \mid x_i > 0\}.$$

Die Spalten von $A_{\mathcal{I}}$ sind nach Satz 6.16 linear unabhängig. Damit muss $|\mathcal{I}| \leqslant m$ gelten. Falls $|\mathcal{I}| < m$ gilt, kann $\mathcal{I}$ (wegen $\operatorname{Rang} A = m$) zu einer Basis B ergänzt werden. (Dann ist x degeneriert.) (2 Punkte)

(b)„⇒"(c) Sei x ein zulässiger Basisvektor mit Basis B. Wir setzen

$$c_j \coloneqq \begin{cases} 0 & \text{falls } j \in B \\ 1 & \text{falls } j \notin B. \end{cases}$$

Dann ist x optimal wegen $c^\top x = 0$ und $c^\top y \ge 0$ für alle zulässigen y. (1 Punkt)

Zudem ist x die einzige Optimallösung, denn für jedes optimale y gilt $c^\top y = 0$ und somit

$$y_j = 0 = x_j \text{ für } j \notin B.$$

Aus $A_B y_B = Ay = b$ folgt dann, dass

$$y_B = A_B^{-1} b = x_B$$

und damit $x = y$. (1 Punkt)

(c)„⇒"(a) Es seien x und c aus $\mathbb{R}^n$ wie gefordert gegeben. Angenommen x wäre kein Extremalpunkt, dann gäbe es $v, w \in P$ und $\lambda \in (0, 1)$, so dass $x = \lambda v + (1 - \lambda) w$. Wir könnten dann abschätzen, dass

$$c^\top x = c^\top(\lambda v + (1-\lambda)w) = \lambda c^\top v + (1-\lambda)c^\top w > \lambda c^\top x + (1-\lambda)c^\top x = c^\top x,$$

was ein Widerspruch ist. (1 Punkt)

Hausaufgabe 5.2 ("Steilster Abstieg" in der Indexwahl im Simplex-Algorithmus) 4 Punkte

Der Simplex-Algorithmus (Algorithmus 7.6) lässt einige Freiheit in der Wahl der Indizes, welche in die Basis aufgenommen werden sollen (Zeile 6) bzw. in die Nichtbasis aufgenommen werden sollen (Zeile 11). Für den Index, der in die Basis aufgenommen werden soll, ist eine Möglichkeit bspw. die Regel des steilsten Abstiegs, welche

$$r \coloneqq \min(\arg\min\{\widetilde{c}_i \mid i \in N\}) \tag{$*$}$$

verwendet.

Geben Sie ein Beispiel an, das zeigt, dass die Wahl $(*)$ nicht zum bestmöglichen Abstieg im Kostenfunktionalwert führen muss.

Lösung.

Die Steilste-Abstieg-Regel nutzt keine Informationen dazu, wie weit der nächste Schritt gehen wird. Wir bauen also ein Beispiel, bei dem die weitere Schrittweite in die Richtung des weniger steilen Abstiegs den Funktionswert insgesamt stärker verbessert.

Wir betrachten dafür das Problem

$$\begin{aligned} \text{Minimiere} \quad & (-1, -2, 0) \begin{pmatrix} x \\ s \end{pmatrix} \quad \text{über } (x, s) \in \mathbb{R}^{2+1} \\ \text{sodass} \quad & x_1 + 4\, x_2 + s = 4 \\ \text{und} \quad & (x, s) \geqslant 0 \end{aligned}$$

für

$$B = \{3\}, \quad N = \{1, 2\}, \quad (x, s) = (0, 0, 4), \quad c^\top \begin{pmatrix} x \\ s \end{pmatrix} = 0.$$

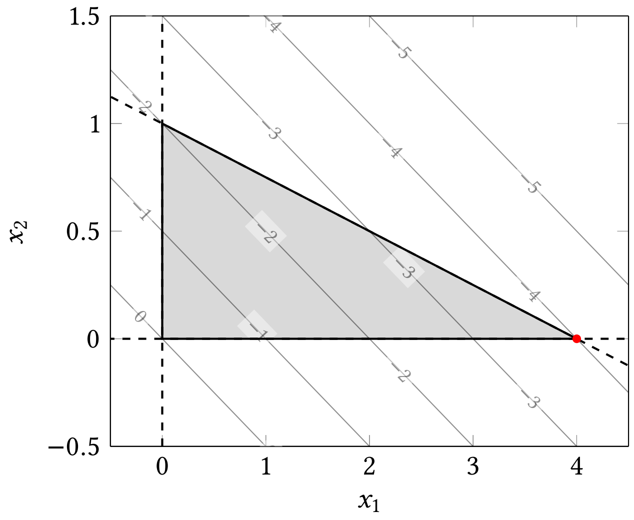
Dann ergeben sich die reduzierten Kosten zu

$$\widetilde{c}_N = c_N - A_N^\top A_B^{-\top} \underbrace{c_B}_{=0} = c_N = \begin{pmatrix} -1 \\ -2 \end{pmatrix}$$

und die Regel $(*)$ liefert den Tausch von $r = 2$ gegen $\ell = 3$, was uns an den Punkt $(x, s) = (0, 1, 0)$ mit $c^\top \begin{pmatrix} x \\ s \end{pmatrix} = -2$ bringt.

Der Tausch $r = 1$ gegen $\ell = 3$ hingegen hätte uns an den Punkt $x = (4, 0, 0)$ mit $c^\top \begin{pmatrix} x \\ s \end{pmatrix} = -4$ und damit direkt an den Minimierer gebracht. (4 Punkte)

Beachte: Das Problem ist die Normalform der Optimierung über ein verzerrtes Standardsimplex in R^2, bei dem der Kostenvektor genau so gewählt ist, dass die Funktionswerte in Richtung x_2 stärker fallen aber die andere Richtung auf Grund der Schrittlänge direkt zum Minimierer geführt hätte.



Hausaufgabe 5.3 (Post-Processing der Simplexphase I) 7 Punkte

Gegeben sei ein lineares Programm in Normalform (Gleichung (6.6)), für das o. B. d. A. $b \geq 0$ ist. Um das Optimierungsproblem mit dem Simplex-Algorithmus (Algorithmus 7.6) lösen zu können, benötigt man eine Anfangsbasis zu einem zulässigen Basisvektor. Wenn $\operatorname{Rang} A = m$ ist und das Problem zulässig ist, dann kann ein solcher zulässiger Basisvektor durch Lösen des linearen Hilfsproblems (**Phase-I-Problem**)

$$\left.\begin{aligned} \text{Minimiere} \quad & \mathbf{1}^\mathsf{T} z \quad \text{über } (x, z) \in \mathbb{R}^n \times \mathbb{R}^m \\ \text{sodass} \quad & Ax + z = b \\ \text{und} \quad & x \geqslant 0, \quad z \geqslant 0 \end{aligned}\right\} \tag{Gleichung (7.7)}$$

mit $\mathbf{1} = (1, 1, \dots, 1)^\mathsf{T} \in \mathbb{R}^m$ mit dem Simplex-Algorithmus bestimmt werden, siehe Satz 7.10. Ist der berechnete optimale Basisvektor $(x^*, 0)^\mathsf{T}$ des Hilfsproblems nicht entartet, dann kann die dazugehörige Basis B^* sofort als Anfangsbasis für die Lösung des Ursprungsproblems verwendet werden. Andernfalls kann die Basis B^* noch Indizes in $\{n+1, \dots, n+m\}$ enthalten und muss daher modifiziert werden. Wie das geht, sehen wir hier.

Es seien $\operatorname{Rang} A = m$ und ein entarteter optimaler Basisvektor des Phase-I-Problems $(x^*, z^*)^\mathsf{T}$ mit $z^* = 0$ zur Basis B^* gegeben. Weiter sei $\ell \in \{n+1, \dots, n+m\} \cap B^*$.

(a) Zeigen Sie, dass ein Index $r \in \{1, \dots, n\} \setminus B^*$ existiert, sodass für den durch $[A, \mathrm{Id}_m]_{B^*} d_{B^*} = a_r$ definierten Vektor d die Aussage $d_\ell \neq 0$ gilt.

(b) Zeigen Sie, dass $(x^*, 0)^\intercal$ für jeden Index r, der die Eigenschaft aus Punkt (a) besitzt, ebenfalls ein Basisvektor zur Indexmenge $B^+ \coloneqq (B^* \cup \{r\}) \setminus \{\ell\}$ ist.

(c) Beschreiben Sie, wie Sie einen Index r, der die Eigenschaft aus Punkt (a) besitzt, praktisch bestimmen können und wie Sie mit diesem Vorgehen aus einem optimalen Basis nach der Phase-I-Optimierung eine Anfangsbasis für die Optimierung des Ursprungsproblems generieren können.

Lösung.

(a) Angenommen kein solch ein $r \in \{1, \dots, n\} \setminus B^*$ existiere, d. h.

$$\text{für alle} \quad r \in \{1, \dots, n\} \setminus B^* \quad \text{mit} \quad [A, \mathrm{Id}_m]_{B^*}\, d_{B^*} = a_r \quad \text{gilt} \quad d_\ell = 0.$$

Dann ist

$$a_r \in \operatorname{span}\{[A, \mathrm{Id}_m]_i\}_{i \in B^* \setminus \{\ell\}} \qquad \forall r \in \{1, \dots, n\} \setminus B^*.$$

Die Nichtbasis-Spalten von A sind also durch $m - 1$ linear unabhängige Spalten darstellbar, die auch die Basis-Spalten von A enthalten. Jede Spalte von A ist also durch $m-1$ linear unabhängige Spalten darstellbar, das Bild von A also höchstens $(m-1)$-dimensional, also $\operatorname{Rang} A \le m - 1$. Das ist ein Widerspruch zur Rangvoraussetzung an A. (2 Punkte)

(b) Wir müssen zeigen, dass die Submatrix $[A, \mathrm{Id}_m]_{B^+}$ regulär ist und

$$[A, \mathrm{Id}_m]_{B^+} \begin{pmatrix} x^* \\ z^* \end{pmatrix}_{B^+} = b.$$

Wegen $(x^*, z^*)_r = (x^*, z^*)_\ell = 0$ gilt offensichtlich

$$[A, \mathrm{Id}_m]_{B^+} \begin{pmatrix} x^* \\ z^* \end{pmatrix}_{B^+} = [A, \mathrm{Id}_m]_{B^*} \begin{pmatrix} x^* \\ z^* \end{pmatrix}_{B^*} = [A, \mathrm{Id}_m] \begin{pmatrix} x^* \\ z^* \end{pmatrix} = b$$

(1 Punkt)

Außerdem ist $[A, \mathrm{Id}_m]_{B^+}$ regulär, denn anderenfalls könnte die Spalte a_r durch die verbleibenden Spalten aus $B^+ \setminus \{r\} = B^* \setminus \{\ell\}$ dargestellt werden können, im Widerspruch zu $d_\ell \neq 0$. Konkret würden α_i, $i \in B^+$ existieren, die nicht alle 0 sind, so dass

$$\begin{aligned}
0 &= \sum_{i \in B^+} \alpha_i a_i \\
&= \sum_{i \in B^+ \setminus \{r\}} \alpha_i a_i + \alpha_r a_r \\
&= \sum_{i \in B^+ \setminus \{r\}} \alpha_i a_i + \alpha_r \sum_{i \in B^* \setminus \{\ell\}} d_i a_i + \alpha_r d_\ell a_\ell \quad (\text{denn } [A, \mathrm{Id}_m]_{B^*} \, d_{B^*} = a_r) \\
&= \sum_{i \in B^* \setminus \{\ell\}} (\alpha_i + \alpha_r d_i) \, a_i + \alpha_r d_\ell a_\ell \quad (\text{denn } B^+ \setminus \{r\} = B^* \setminus \{\ell\}),
\end{aligned}$$

weshalb wegen $d_\ell \neq 0$ auch $[A, \mathrm{Id}_m]_{B^*}$ nicht regulär wäre, im Widerspruch zur Basiseigenschaft von B^*. (2 Punkte)

(c) Für einen Index $\ell \in \{n+1, \ldots, n+m\} \cap B^*$ bestimmen wir $[A, \mathrm{Id}_m]_{B^*}^{-1} \, [A]_{\{1,\ldots,n\} \setminus B^*}$. In der zu ℓ-gehörigen Zeile finden wir einen Nicht-Null-Eintrag, dessen Spaltenindex wir gegen ℓ in die Basis reintauschen. (1 Punkt)

Das wiederholen wir, bis alle Basisindizes in $\{1, \ldots, n\}$ liegen. (1 Punkt)

Zusatzaufgabe 5.4 (Zyklen im Simplex-Verfahren haben mindestens die Länge 3) 10 Bonuspunkte

Zeigen Sie, dass ein im Simplex-Verfahren soeben aus der Basis entfernter Index im nächsten Simplex-Schritt nicht sofort wieder in die Basis aufgenommen werden wird.

Hinweis: Zeigen Sie, dass der Vektor der reduzierten Kosten im betreffenden Eintrag positiv sein wird. Verwenden Sie dazu die Cramersche Regel.

Lösung.

Wir befinden uns im Simplex-Verfahren (Algorithmus 7.6) in dem Schritt, der den Index ℓ aus der Basis entfernt und dafür r aufnimmt. Hier passieren mit der aktuellen Basis B, die ℓ noch enthält, und der Nichtbasis N also folgende Schritte:

(a) Berechne die reduzierten Kosten $\widetilde{c}_N \coloneqq c_N - A_N^\top A_B^{-\top} c_B$

(b) Wähle einen Index $r \in N$ mit $\widetilde{c}_r < 0$

(c) Berechne $\Delta x_B \coloneqq -A_B^{-1} \, a_r$

(d) Bestimme $\hat{t} \geqslant 0$ und $\ell \in B$ gemäß $\hat{t} \coloneqq \min\left\{ -\frac{x_i}{\Delta x_i} \,\middle|\, i \in B,\ \Delta x_i < 0 \right\} = -\frac{x_\ell}{\Delta x_\ell}$

(e) Setze

$$x_i^+ \coloneqq \begin{cases} x_i + \hat{t}\,\Delta x_i & \text{für } i \in B,\ i \neq \ell, \\ \hat{t} & \text{für } i = r, \\ 0 & \text{sonst} \end{cases}$$

Setze $B^+ \coloneqq \big(B \cup \{r\}\big) \setminus \{\ell\}$
Setze $N^+ \coloneqq \big(N \cup \{\ell\}\big) \setminus \{r\}$

(f) Berechne die neuen reduzierten Kosten

$$\widetilde{c}_N^+ \coloneqq c_{N^+} - A_{N^+}^\top A_{B^+}^{-\top} c_{B^+}$$

Wir zeigen, dass $\widetilde{c}_\ell^+ > 0$ ist, so dass ℓ im folgenden pricing-Schritt nicht als neue Basisvariable in Frage kommt, bzw. der Algorithmus in einem Optimierer terminiert. Dabei ist

$$\widetilde{c}_\ell^+ = c_\ell - a_\ell^\top A_{B^+}^{-\top} c_{B^+}.$$

(1 Punkt)

Abschätzungsinformation zu dem ersten Term c_ℓ erhalten wir aus den Vorzeicheninformationen zu $\widetilde{c}_r$ aus Punkt (b), weil

$$0 > \widetilde{c}_r = c_r - a_r^\top A_B^{-\top} c_B = c_r + \Delta x_B^\top c_B = c_r + \sum_{i \in B} \Delta x_i c_i = c_r + \sum_{i \in B \setminus \{\ell\}} \Delta x_i c_i + c_\ell \Delta x_\ell$$

(2 Punkte)

ist, und damit wegen $\Delta x_\ell < 0$ (Punkt (d)) auch

$$c_\ell > -\frac{1}{\Delta x_\ell}\left(c_r + \sum_{i \in B \setminus \{\ell\}} \Delta x_i c_i\right).$$

(1 Punkt)

Es gilt also

$$\begin{aligned} \widetilde{c}_\ell^+ &= c_\ell - a_\ell^\top A_{B^+}^{-\top} c_{B^+} \\ &= c_\ell \underbrace{- \big(A_{B^+}^{-1} a_\ell\big)}_{=:\Delta x_{B^+}^+}{}^\top c_{B^+} \\ &> -\frac{1}{\Delta x_\ell}\left(c_r + \sum_{i \in B \setminus \{\ell\}} \Delta x_i c_i\right) + \Delta x_{B^+}^{+\top} c_{B^+}. \end{aligned}$$

Dass die rechte Seite 0 ist, werden wir nun mit der Cramerschen Regel zeigen.

Beachte: Bei der Anwendung der Cramerschen Regel ist wichtig zu erkennen, dass die Reihenfolge der Spalten in der Matrix keine entscheidende Rolle spielt – durch Umsortieren der Lösungseinträge erhält man auch immer die Lösung zu dem Gleichungssystem mit der umsortierten Matrix. In der Cramerschen Regel sieht man das daran, dass wir durch die zwei Determinanten teilen. Ein Umsortieren der Spalten passiert also sowohl in der Matrix im Zähler, wie auch in der im Nenner. Permutation der Spalten ändert im Zweifel also das Vorzeichen der Determinanten, aber immer in beiden, wodurch sich der Lösungseintrag nicht ändert. Für ein sauberes Aufschreiben unserer Lösung macht es aber Sinn, sich auf eine konkrete Reihenfolge einzulassen, da die Cramersche Regel auf verschiedene Matrizen angewandt wird, die sich jeweils nur in der Spaltenreihenfolge unterscheiden. Die Reihenfolge spielt beim Vergleich der Vorzeichen der Determinanten dieser Matrizen dann sehr wohl eine Rolle. Zwei offensichtliche Möglichkeiten bieten sich an: Entweder aufsteigende Sortierung der Indizes oder die Urpsrungsordnung der Basis B beibehaltend. Da Variante 1 zu störenden Termen in den Gleichungen führt, werden wir Variante zwei wählen.

Wir bezeichnen mit

$$
\begin{aligned}
B &= \{i_1, i_2, \dots, i_{k_0-1}, \underbrace{i_{k_0}}_{=\ell}, i_{k_0+1}, \dots, i_m\} = \{i_1, i_2, \dots, i_{k_0-1}, \ell, i_{k_0+1}, \dots, i_m\} \\
B^+ &= \{i_1, i_2, \dots, i_{k_0-1}, \underbrace{i_{k_0}}_{=r}, i_{k_0+1}, \dots, i_m\} = \{i_1, i_2, \dots, i_{k_0-1}, r, i_{k_0+1}, \dots, i_m\}
\end{aligned}
$$

die endlichen Folgen der Indexelemente aus den Indexmengen B, B^+ bezüglich einer durch B festgelegten Reihenfolge (es wird also nicht nur ℓ durch r getauscht, sondern die restliche Reihenfolge beibehalten). Weiterhin bezeichnen wir für $k \neq k_0$ die entsprechenden Folgen zu den modifizierten Indexmengen mit

$$
\begin{aligned}
B \setminus \{i_k\} \cup \{r\} &= \{i_1, i_2, \dots, i_{k-1}, r, i_{k+1}, \dots, i_{k_0-1}, \underbrace{i_{k_0}}_{=\ell}, i_{k_0+1}, \dots, i_m\} \\
B^+ \setminus \{i_k\} \cup \{\ell\} &= \{i_1, i_2, \dots, i_{k-1}, \ell, i_{k+1}, \dots, i_{k_0-1}, \underbrace{i_{k_0}}_{=r}, i_{k_0+1}, \dots, i_m\}
\end{aligned}
$$

und entsprechend die dazugehörigen Matrizen mit den dazugehörig ausgewählten Spalten mit

$$
\begin{aligned}
A_B &= (a_{i_1}, a_{i_2}, \dots, a_{i_{k-1}}, a_\ell, a_{i_{k+1}}, \dots, i_m) \\
A_{B^+} &= (a_{i_1}, a_{i_2}, \dots, a_{i_{k-1}}, a_r, a_{i_{k+1}}, \dots, i_m)
\end{aligned}
$$

und wieder für $k \neq k_0$ mit

$$\begin{aligned}
A_{B\setminus\{i_k\}\cup\{r\}} &= \{a_{i_1}, a_{i_2}, \dots, a_{i_{k-1}}, a_r, a_{i_{k+1}}, \dots, a_{i_{k_0-1}}, \underbrace{a_{i_{k_0}}}_{=a_\ell}, a_{i_{k_0+1}}, \dots, a_{i_m}\} \\
A_{B^+\setminus\{i_k\}\cup\{\ell\}} &= \{a_{i_1}, a_{i_2}, \dots, a_{i_{k-1}}, a_\ell, a_{i_{k+1}}, \dots, a_{i_{k_0-1}}, \underbrace{a_{i_{k_0}}}_{=a_r}, a_{i_{k_0+1}}, \dots, a_{i_m}\}
\end{aligned}$$

Die Cramersche Regel liefert sofort den Zusammenhang

$$\Delta x_\ell = -\frac{\det A_{B^+}}{\det A_B} = \frac{1}{\Delta x_r^+}$$

(2 Punkte)

sowie für alle $i_k \in B$ mit $k \neq k_0$ (oder in Mengenschreibweise $i_k \in B \setminus \{\ell\} = B^+ \setminus \{r\}$) den Zusammenhang

$$\Delta x_{i_k} = -\frac{\det A_{B\setminus\{i_k\}\cup\{r\}}}{\det A_B} \quad \text{und} \quad \Delta x_{i_k}^+ = -\frac{\det A_{B^+\setminus\{i_k\}\cup\{\ell\}}}{\det A_{B^+}}$$

wobei auf Grund der Reihenfolge der Spalten a_r, a_ℓ in $A_{B\setminus\{i_k\}\cup\{r\}}$ und $A_{B^+\setminus\{i_k\}\cup\{\ell\}}$ gilt, dass

$$\det A_{B\setminus\{i_k\}\cup\{r\}} = -\det A_{B^+\setminus\{i_k\}\cup\{\ell\}},$$

denn eine der jeweiligen Matrizen lässt sich durch eine ungerade Anzahl von Spaltenpermutationen in die jeweils andere überführen, was das Vorzeichen der Determinante flipt. (2 Punkte)

Damit können wir weiter abschätzen:

$$\begin{aligned}
\widetilde{c}_\ell^+ &> -\frac{1}{\Delta x_\ell}\left(c_r + \sum_{i\in B\setminus\{\ell\}} \Delta x_i c_i\right) + \Delta x_{B^+}^{+\top} c_{B^+} \\
&= -\Delta x_r^+ c_r - \sum_{i\in B\setminus\{\ell\}} \Delta x_r^+ \Delta x_i c_i + \Delta x_r^+ c_r + \sum_{\underbrace{i\in B^+ \setminus \{r\}}_{=B\setminus\{\ell\}}} \Delta x_i^+ c_i \\
&= \sum_{i\in B\setminus\{\ell\}} \left(-\Delta x_r^+ \Delta x_i + \Delta x_i^+\right) c_i \\
&= \sum_{i\in B\setminus\{\ell\}} c_i \, \frac{\overbrace{\det A_{B\setminus\{i\}\cup\{r\}} + \det A_{B^+\setminus\{i\}\cup\{\ell\}}}^{=0}}{\det A_{B^+}} \\
&= 0.
\end{aligned}$$

(2 Punkte)

Zusatzaufgabe 5.5 (Implementierung eines Simplexverfahrens) 15 Bonuspunkte

Bearbeiten Sie `P2_Simplexverfahren.ipynb`.

Lösung.

Siehe `P2_Simplexverfahren_solutions.ipynb`. (15 Punkte)

Bitte reichen Sie Ihre Lösungen als ein PDF auf Mampf ein.