11 SQP–Verfahren

In den letzten Jahren haben sich die sogenannten SQP–Verfahren (Verfahren der **s**equentiellen **q**uadratischen **P**rogrammierung) als Standardverfahren zur Lösung allgemeiner restringierter Optimierungsaufgaben durchgesetzt. Die Verfahren dieser Klasse gehen zurück auf Arbeiten von Biggs (1975), Han (1977) und Powell (1978). Ausgefeilte Realisierungen von SQP–Verfahren stammen von Gill et al. (1984) und Schittkowski (1985). Herausragende Eigenschaften dieser Verfahren sind ihre schnelle Konvergenz (die Verfahren beruhen i.w. auf einem Newton–Verfahren) und eine geringe Anzahl benötigter Funktionsauswertungen im Vergleich zu anderen Verfahren.

Realisierungen von SQP–Verfahren findet man auch in den Programmbibliotheken NAG und IMSL und in der Optimierungs–Toolbox von MATLAB (Algorithmus **constr**).

Das Grundprinzip der SQP–Verfahren beruht in einer schrittweisen Ersetzung des Ausgangsproblems durch ein quadratisches Hilfsproblem (d.h. durch eine quadratische Approximation des Zielfunktionals und eine linearen Approximation der Gleichungs- und Ungleichungsnebenbedingungen). Das quadratische Hilfsproblem läßt sich mit den im letzten Abschnitt besprochenen Techniken (Reduktion und Strategien zur Bestimmung der aktiven Menge) effizient lösen.

A. Die Lagrange–Newton–Iteration

Ausgangspunkt für die SQP–Verfahren ist die sogenannte Lagrange–Newton–Iteration für Optimierungsaufgaben mit Gleichungsrestriktionen. Wir betrachten also zunächst die Aufgabe

Minimiere eine (glatte) Funktion

$$f(\boldsymbol{x}), \;\; \boldsymbol{x} \in \mathbb{R}^n \tag{11.1}$$

unter den (glatten) Nebenbedingungen

$$c_i(\boldsymbol{x}) \;=\; 0, \qquad i = 1, \ldots, m \tag{11.2}$$

Ist $\boldsymbol{x}^*$ ein reguläres Minimum von f unter der Nebenbedingung (11.2), so gilt mit der Lagrange–Funktion

$$L(\boldsymbol{x}, \boldsymbol{\lambda}) \;:=\; f(\boldsymbol{x}) \;-\; \sum_{i=1}^{m} \lambda_i \, c_i(\boldsymbol{x}) \tag{11.3}$$

($\boldsymbol{\lambda}$: Lagrange–Parameter) nach Satz (7.5) die folgende notwendige Bedingung:

$$\nabla L(\boldsymbol{x}, \boldsymbol{\lambda}) \;:=\; \begin{pmatrix} \boldsymbol{\nabla} f(\boldsymbol{x}) \;-\; \sum\limits_{i=1}^{m} \lambda_i \, \boldsymbol{\nabla} c_i(\boldsymbol{x}) \\ -\boldsymbol{c}(\boldsymbol{x}) \end{pmatrix} \;=\; \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix} . \tag{11.4}$$

Man beachte, daß die Gradientenbildung der Lagrange-Funktion bezüglich aller Variablen $(\boldsymbol{x}, \boldsymbol{\lambda})$ erfolgt. $\boldsymbol{c}$ ist wie früher erklärt $\boldsymbol{c}(\boldsymbol{x}) := (c_1(\boldsymbol{x}), \ldots, c_m(\boldsymbol{x}))^{\mathrm{T}}$.

Die Kuhn–Tucker–Bedingungen (11.4) bilden ein nichtlineares Gleichungssystem zur Berechnung von $(\boldsymbol{x}, \boldsymbol{\lambda})$. Hierzu läßt sich nun das Newton–Verfahren angewenden. Nach

(7.8) erhält man das folgende lineare Gleichungssystem (Newton–Gleichung) für die Newton–Korrekturen $(\boldsymbol{\Delta x}, \boldsymbol{\Delta\lambda})$:

$$\boldsymbol{\nabla}^2 L(\boldsymbol{x}, \boldsymbol{\lambda}) \begin{pmatrix} \boldsymbol{\Delta x} \\ \boldsymbol{\Delta\lambda} \end{pmatrix} = -\boldsymbol{\nabla} L(\boldsymbol{x}, \boldsymbol{\lambda}) \ . \tag{11.5}$$

Die Hesse–Matrix der Lagrange–Funktion hat dabei die folgende Struktur

$$\boldsymbol{\nabla}^2 L(\boldsymbol{x}, \boldsymbol{\lambda}) = \begin{pmatrix} \boldsymbol{W} & -\boldsymbol{A}^{\mathrm{T}} \\ -\boldsymbol{A} & \boldsymbol{0} \end{pmatrix} \in \mathbb{R}^{(n+m,n+m)} \qquad \text{mit} \tag{11.6}$$

$$\boldsymbol{W} = \boldsymbol{\nabla}^2 f(\boldsymbol{x}) - \sum_{i=1}^{m} \lambda_i \, \boldsymbol{\nabla}^2 c_i(\boldsymbol{x}) \tag{11.7}$$

$$\boldsymbol{A} = \boldsymbol{c}'(\boldsymbol{x}) \qquad \text{(Jacobi–Matrix)} \ . \tag{11.8}$$

Setzt man (11.6)–(11.8) in die Newton–Gleichung ein, so erhält man mit $\boldsymbol{\lambda}^{\mathrm{neu}} := \boldsymbol{\lambda} + \boldsymbol{\Delta\lambda}$, $\boldsymbol{g} := \boldsymbol{\nabla} f(\boldsymbol{x})$ und $\boldsymbol{c} := \boldsymbol{c}(\boldsymbol{x})$

$$\begin{aligned} \begin{pmatrix} \boldsymbol{W} & -\boldsymbol{A}^{\mathrm{T}} \\ -\boldsymbol{A} & \boldsymbol{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{\Delta x} \\ \boldsymbol{\lambda}^{\mathrm{neu}} \end{pmatrix} &= \begin{pmatrix} -\boldsymbol{g} \\ \boldsymbol{c} \end{pmatrix} \\ \boldsymbol{x}^{\mathrm{neu}} &= \boldsymbol{x} + \boldsymbol{\Delta x} \ . \end{aligned} \tag{11.9}$$

Einige theoretische Eigenschaften der Lagrange–Newton–Iteration (11.5) bzw. (11.9) sind im folgenden Satz zusammengefaßt

Satz (11.10) *Ist $\boldsymbol{x}^*$ ein reguläres lokales Minimum von f auf $M := \boldsymbol{c}^{-1}(\boldsymbol{0})$ mit Lagrange–Multiplikatoren $\boldsymbol{\lambda}^*$ und sind die hinreichenden Bedingungen (7.9) erfüllt. d.h. $\boldsymbol{W}$ ist positiv definit auf dem Tangentialraum $T_{\boldsymbol{x}^*}M = \{\boldsymbol{y} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{y} = \boldsymbol{0}\}$, so gelten folgende Eigenschaften:*

a) Die Hesse–Matrix $\boldsymbol{\nabla}^2 L(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$ ist regulär.

b) Die Lagrange–Newton–Iteration konvergiert für alle Startwerte $(\boldsymbol{x}^0, \boldsymbol{\lambda}^0)$ mit $\|\boldsymbol{x}^0 - \boldsymbol{x}^*\|$ hinreichend klein und $\boldsymbol{\nabla}^2 L(\boldsymbol{x}^0, \boldsymbol{\lambda}^0)$ regulär. Die Konvergenz ist quadratisch.

B. SQP-Verfahren für gleichungsrestringierte Optimierungsaufgaben

Im Fall einer unrestringierten Optimierungsaufgabe: Minim. $f(\boldsymbol{x})$, $\boldsymbol{x} \in \mathbb{R}^n$, läßt sich die Newton–Korrektur $\boldsymbol{\Delta x}$ mit $\boldsymbol{\nabla}^2 f(\boldsymbol{x}) \, \boldsymbol{\Delta x} = -\boldsymbol{\nabla} f(\boldsymbol{x})$ offensichtlich als Lösung der folgenden quadratischen Optimierungsaufgabe erhalten

$$\text{Minimiere} \quad q(\boldsymbol{\Delta x}) := \frac{1}{2} \boldsymbol{\Delta x}^{\mathrm{T}} \, \boldsymbol{\nabla}^2 f(\boldsymbol{x}) \, \boldsymbol{\Delta x} + \boldsymbol{\nabla} f(\boldsymbol{x})^{\mathrm{T}} \, \boldsymbol{\Delta x} + f(\boldsymbol{x}) \, .$$

Hierbei muß die Hesse–Matrix $\boldsymbol{\nabla}^2 f(\boldsymbol{x})$ als positiv definit vorausgesetzt werden.

Eine analoge Eigenschaft läßt sich nun auch für gleichungsrestringierte Optimierungsaufgaben feststellen.

Satz (11.11) *Ist die Hesse–Matrix $\boldsymbol{W}$ für vorgegebene $(\boldsymbol{x}, \boldsymbol{\lambda})$ positiv definit auf dem Tangentialraum $T_{\boldsymbol{x}}M$, so läßt sich die Newton–Korrektur $\boldsymbol{\Delta x}$ aus (11.9) auch als Lösung der folgenden quadratischen Optimierungsaufgabe mit linearen Gleichungsrestriktionen erhalten:*

$$\begin{aligned} &\text{Minimiere} \quad q(\boldsymbol{\Delta x}) \;:=\; \frac{1}{2}\,\boldsymbol{\Delta x}^{\mathrm{T}}\,\boldsymbol{W}\,\boldsymbol{\Delta x} \;+\; \boldsymbol{g}^{\mathrm{T}}\,\boldsymbol{\Delta x} \;+\; f \\ &\text{unter den Nebenbedingungen:} \quad \boldsymbol{A}\,\boldsymbol{\Delta x} \;+\; \boldsymbol{c} \;=\; \boldsymbol{0}\;. \end{aligned} \tag{11.12}$$

$\boldsymbol{W}$ und $\boldsymbol{A}$ sind hierbei durch (11.7) und (11.8) gegeben. Ferner ist wie zuvor $\boldsymbol{g} := \boldsymbol{\nabla} f(\boldsymbol{x})$, $f := f(\boldsymbol{x})$ und $\boldsymbol{c} := \boldsymbol{c}(\boldsymbol{x})$. Der Lagrange–Multiplikator $\boldsymbol{\lambda}^{neu}$ ist zugleich der Lagrange–Multiplikator für die restringierte Optimierungsaufgabe (11.12).

Beweis:

Erfüllt $(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$ die hinreichenden Bedingungen für ein lokales Minimum von f auf M und ist $\boldsymbol{x}^*$ ein regulärer Punkt, so ist das quadratische Hilfsproblem (11.12) für $(\boldsymbol{x}, \boldsymbol{\lambda})$ in einer Umgebung von $(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$ eindeutig lösbar. Für solche $(\boldsymbol{x}, \boldsymbol{\lambda})$ lautet die erweiterte Funktion zu (11.12)

$$Q(\boldsymbol{x}) \;=\; \frac{1}{2}\,\boldsymbol{\Delta x}^{\mathrm{T}}\,\boldsymbol{W}\,\boldsymbol{\Delta x} \;+\; \boldsymbol{g}^{\mathrm{T}}\,\boldsymbol{\Delta x} \;+\; f \;-\; \boldsymbol{\mu}^{\mathrm{T}}\,(\boldsymbol{A}\,\boldsymbol{\Delta x} \;+\; \boldsymbol{c})$$

mit einem Lagrange–Parameter $\boldsymbol{\mu}$. Die notwendige Bedingung erster Ordnung lautet dann

$$\boldsymbol{\nabla} Q(\boldsymbol{x}) \;=\; \boldsymbol{W}\,\boldsymbol{\Delta x} + \boldsymbol{g} \;-\; \boldsymbol{A}^{\mathrm{T}}\,\boldsymbol{\mu} \;=\; \boldsymbol{0}\,.$$

Zusammen mit der Nebenbedingung von (11.12) ist dies genau die Lagrange–Newton–Gleichung (11.9) mit $\boldsymbol{\mu} = \boldsymbol{\lambda}^{\mathrm{neu}}$. □

Bemerkung

Das Hilfsproblem (11.12) kann man sich durch Taylor–Entwicklung von $f(\boldsymbol{x} + \boldsymbol{\Delta x})$ und $\boldsymbol{c}(\boldsymbol{x} + \boldsymbol{\Delta x})$ um $\boldsymbol{\Delta x} = \boldsymbol{0}$ entstanden denken. Es ist jedoch zu beachten, daß im quadratischen Teil der Zielfunktion die Hesse–Matrix der Lagrange–Funktion (und nicht die von f) auftritt.

Zusammengefaßt läßt sich das SQP–Verfahren für Optimierungsprobleme mit Gleichungsrestriktionen nun folgendermaßen formulieren.

Algorithmus (SQP für Gleichungsrestriktionen) (11.13)

(i) $\boldsymbol{x}^0 \in \mathbb{R}^n$, $\boldsymbol{\lambda}^0 \in \mathbb{R}^m$: Startvektoren

(ii) Für $k = 0, 1, \ldots$:

1.) Man setze $f := f(\boldsymbol{x}^k)$, $\boldsymbol{c} := \boldsymbol{c}(\boldsymbol{x}^k)$, $\boldsymbol{g} := \boldsymbol{\nabla} f(\boldsymbol{x}^k)$, $\boldsymbol{A} := \boldsymbol{c}'(\boldsymbol{x}^k)$ und $\boldsymbol{W} := \boldsymbol{\nabla}^2 f(\boldsymbol{x}^k) \;-\; \sum\limits_{i=1}^{m} \lambda_i^k\,\boldsymbol{\nabla}^2 c_i(\boldsymbol{x}^k)$.

2.) Lösung des quadratischen Hilfsproblems (11.12).

3.) Bestimmung von $\boldsymbol{\lambda}^{k+1}$ aus $\boldsymbol{A}^{\mathrm{T}}\,\boldsymbol{\lambda}^{k+1} \;=\; \boldsymbol{W}\,\boldsymbol{\Delta x} \;+\; \boldsymbol{g}$

4.) $\boldsymbol{x}^{k+1} \;:=\; \boldsymbol{x}^k + \boldsymbol{\Delta x}$

5.) Abbruch, falls $\|\boldsymbol{\Delta x}\|$ hinreichend klein ist.

Bemerkungen (11.14)

a) Im Fall Rang$(\boldsymbol{A}) = m$ und $\boldsymbol{W}$ positiv definit auf Kern$(\boldsymbol{A})$ ist die Existenz eines Lagrange–Parameters für das quadratische Hilfsproblem und damit auch die Existenz von $\boldsymbol{\lambda}^{k+1}$ gesichert.

b) Das in Schritt 3.) des Algorithmus angegebene lineare Gleichungssystem für $\boldsymbol{\lambda}^{k+1}$ ist überbestimmt. Es empfiehlt sich daher $\boldsymbol{\lambda}^{k+1}$ als Lösung des linearen Ausgleichsproblems

$$\text{Minimiere} \quad \|\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda} - \boldsymbol{W}\,\boldsymbol{\Delta x} - \boldsymbol{g}\|_2, \;\; \boldsymbol{\lambda} \in \mathbb{R}^m$$

zu ermitteln ($\boldsymbol{Q}\,\boldsymbol{R}$–Zerlegung von $\boldsymbol{A}^{\mathrm{T}}$; vergleiche auch die Reduktionsmethode in Abschnitt 10.1, S. 85)

c) Natürlich ist es sinnvoll, das Newton–Verfahren wie früher durch eine Schrittlängenstrategie zu modifizieren. Hierzu wird der Schritt 4.) ersetzt durch

$$\boldsymbol{x}^{k+1} = \boldsymbol{x}^k + \alpha_k\,\boldsymbol{\Delta x}\,.$$

Die Schrittlänge wird wieder so bestimmt, daß eine bestimmte Testfunktion in jedem Iterationsschritt fällt. Es sind verschiedene Testfunktionen (z.B. nach Han oder Powell) in Gebrauch.

C. Übertragung auf Ungleichungsrestriktionen

Während sich der Lagrange–Newton–Ansatz nicht ohne weiteres auf Ungleichungsnebenbedingungen übertragen läßt, bereiten diese für den SQP–Ansatz keine Probleme. Die Frage, welche dieser Nebenbedingungen jeweils aktiv sind, wird an das quadratische Hilfsproblem weitergegeben.

Für die restringierte Optimierungsaufgabe

$$\begin{array}{lll} \text{Minimiere} & f(\boldsymbol{x}), & \boldsymbol{x} \in \mathbb{R}^n \\ \text{Nebenbedingungen} & c_i(\boldsymbol{x}) = 0, & i \in E := \{1, \ldots, m'\} \\ & c_i(\boldsymbol{x}) \geq 0, & i \in I := \{m'+1, \ldots, m\} \end{array} \tag{11.15}$$

setze man zu vorgegebenen $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{\lambda} \in \mathbb{R}^m$ (wie früher):

$$f := f(\boldsymbol{x}), \quad \boldsymbol{g} := \boldsymbol{\nabla} f(\boldsymbol{x}), \quad \boldsymbol{W} := \boldsymbol{\nabla}^2 f(\boldsymbol{x}) - \sum_{i=1}^m \lambda_i^k\,\boldsymbol{\nabla}^2 c_i(\boldsymbol{x})\,.$$

Das zugehörige **quadratische Hilfsproblem** wird dann folgendermaßen definiert

$$\begin{array}{ll} \text{Minimiere} & q(\boldsymbol{\Delta x}) = \frac{1}{2}\,\boldsymbol{\Delta x}^{\mathrm{T}}\,\boldsymbol{W}\,\boldsymbol{\Delta x} + \boldsymbol{g}^{\mathrm{T}}\,\boldsymbol{\Delta x} + f \\ \text{Nebenbedingungen} & \boldsymbol{\nabla} c_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\Delta x} + c_i(\boldsymbol{x}) = 0, \quad i \in E \\ & \boldsymbol{\nabla} c_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\Delta x} + c_i(\boldsymbol{x}) \geq 0, \quad i \in I\,. \end{array} \tag{11.16}$$

Bei der Lösung dieses quadratischen Hilfsproblems werden mittels der Strategie der aktiven Menge (vergleiche Abschnitt 10.2, S. 90) die **Indizes der aktiven linearisierten Ungleichungsnebenbedingungen**

$$J_{\text{lin}} \;:=\; \{i \in I \;:\; \boldsymbol{\nabla} c_i(\boldsymbol{x})^{\mathrm{T}} \boldsymbol{\Delta x} \;+\; c_i(\boldsymbol{x}) \;=\; 0\}$$

bestimmt. $\boldsymbol{A}'$ bezeichne die Matrix mit den Zeilen $\boldsymbol{\nabla} c_i(\boldsymbol{x})^{\mathrm{T}}$, $i \in J_{\text{lin}} \cup E$.

Ist $\boldsymbol{\Delta x}$ Lösung von (11.16) und hat die zugehörige Matrix $\boldsymbol{A}'$ vollen Rang (d.h. $\boldsymbol{\Delta x}$ ist ein regulärer Punkt), so existieren aufgrund der Kuhn–Tucker–Bedingungen Lagrange–Multiplikatoren λ_i^{neu}, $i \in J_{\text{lin}} \cup E$ mit

$$\boldsymbol{\nabla} Q(\boldsymbol{x}) \;=\; \boldsymbol{W}\,\boldsymbol{\Delta x} \;+\; \boldsymbol{g} \;-\; \boldsymbol{A}'^{\mathrm{T}}\,\boldsymbol{\lambda}^{\text{neu}} \;=\; \boldsymbol{0}$$

und es gilt $\lambda_i^{\text{neu}} \geq 0$ für $i \in J_{\text{lin}}$.

$Q(\boldsymbol{x})$ bezeichnet dabei die erweiterte Zielfunktion zu (11.16).

Insgesamt ist damit der der folgende Algorithmus (analog zu (11.13)) begründet:

Algorithmus (SQP Verfahren) (11.17)

(i) $\boldsymbol{x}^0 \in \mathbb{R}^n$, $\boldsymbol{\lambda}^0 \in \mathbb{R}^m$: Startvektoren

(ii) Für $k = 0, 1, \ldots$:

1.) Man setze $f := f(\boldsymbol{x}^k)$, $\boldsymbol{g} := \boldsymbol{\nabla} f(\boldsymbol{x}^k)$, und $\boldsymbol{W} := \boldsymbol{\nabla}^2 f(\boldsymbol{x}^k) - \sum\limits_{i \in E \cup I} \lambda_i^k \boldsymbol{\nabla}^2 c_i(\boldsymbol{x}^k)$.

2.) Lösung des quadratischen Hilfsproblems (11.12) mittels Reduktion und einer Strategie der aktiven Menge. Bestimmung von $\boldsymbol{\Delta x}$, J_{lin} und $\boldsymbol{A}'$.

3.) Bestimmung von $\boldsymbol{\lambda}^{\text{neu}}$ aus $\boldsymbol{A}'^{\mathrm{T}}\,\boldsymbol{\lambda}^{\text{neu}} \;=\; \boldsymbol{W}\,\boldsymbol{\Delta x} \;+\; \boldsymbol{g}$ (etwa über die Lösung eines linearen Ausgleichsproblems!)

4.) $\boldsymbol{x}^{k+1} \;:=\; \boldsymbol{x}^k + \boldsymbol{\Delta x}$

5.) $\lambda_i^{k+1} \;:= \begin{cases} \lambda_i^{\text{neu}}, & \text{falls } i \in E \cup J_{\text{lin}} \\ 0, & \text{falls } i \in I \setminus J_{\text{lin}} \end{cases}$

6.) Abbruch, falls $\|\boldsymbol{\Delta x}\|$ hinreichend klein ist.

Der Ablauf des SQP–Verfahrens wollen wir an dem folgenden einfachen Beispiel erläutern.

Beispiel (11.18) (nach C. Geiger)

Wir betrachten die restringierte Optimierungsaufgabe

$$\begin{array}{ll} \text{Minimiere} & f(x_1,x_2) \;:=\; -\,x_1 \;-\; x_2 \\ \text{Nebenbedingungen} & c_1(x_1,x_2) \;:=\; x_2 \;-\; x_1^2 \;\geq\; 0 \\ & c_2(x_1,x_2) \;:=\; 1 \;-\; x_1^2 \;-\; x_2^2 \;\geq\; 0\,. \end{array}$$

Eine geometrische Veranschaulichung von Nebenbedingungen und Zielfunktion zeigt, daß das Problem eine eindeutig bestimmte Lösung besitzt, nämlich

$$\boldsymbol{x}^* \;=\; (1/\sqrt{2}, 1/\sqrt{2})^{\mathrm{T}}, \quad J(\boldsymbol{x}^*) \;=\; \{2\}, \quad f(\boldsymbol{x}^*) \;=\; -\sqrt{2}\,.$$

Zur Anwendung des SQP–Verfahrens werden noch folgende Informationen benötigt:

$$f \;=\; -x_1 - x_2, \qquad \boldsymbol{g} \;=\; \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \qquad \boldsymbol{\nabla}^2 f(\boldsymbol{x}) \;=\; \boldsymbol{0}$$

$$c_1 \;=\; x_2 - x_1^2, \qquad \boldsymbol{\nabla} c_1 \;=\; \begin{pmatrix} -2x_1 \\ 1 \end{pmatrix}, \quad \boldsymbol{\nabla}^2 c_1(\boldsymbol{x}) \;=\; \begin{pmatrix} -2 & 0 \\ 0 & 0 \end{pmatrix}$$

$$c_2 \;=\; 1 - x_1^2 - x_2^2, \quad \boldsymbol{\nabla} c_2 \;=\; \begin{pmatrix} -2x_1 \\ -2x_2 \end{pmatrix}, \quad \boldsymbol{\nabla}^2 c_2(\boldsymbol{x}) \;=\; \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}$$

Das quadratische Hilfsproblem ist demnach gegeben durch die Zielfunktion

$$q(\boldsymbol{\Delta x}) \;=\; (\lambda_1 + \lambda_2)\,\Delta x_1^2 \;+\; \lambda_2\,\Delta x_2^2 \;-\; (\Delta x_1 + \Delta x_2) \;-\; (x_1 + x_2)$$

und die Nebenbedingungen

$$\begin{aligned} -2\,x_1\,\Delta x_1 \;+\; \Delta x_2 \;+\; (x_2 - x_1^2) \;&\geq\; 0 \\ -2\,x_1\,\Delta x_1 \;-\; 2\,x_2\,\Delta x_2 \;+\; (1 - x_1^2 - x_2^2) \;&\geq\; 0\,. \end{aligned}$$

Wir starten das Verfahren mit den Anfangsnäherungen $x_1^0 \;=\; 1/2,\;\; x_2^0 \;=\; 1$ und $\lambda_1^0 \;=\; \lambda_2^0 \;=\; 0$. Im ersten Iterationsschritt ist dann die folgende lineare Optimierungsaufgabe zu lösen:

$$\begin{array}{ll} \text{Minimiere} & q(\boldsymbol{\Delta x}) \;=\; -\,\dfrac{3}{2} \;-\; \Delta x_1 \;-\; \Delta x_2 \\[2ex] \text{Nebenbedingungen} & -\Delta x_1 \;+\; \Delta x_2 \;+\; \dfrac{3}{4} \;\geq\; 0 \\[2ex] & -\Delta x_1 \;-\; 2\Delta x_2 \;-\; \dfrac{1}{4} \;\geq\; 0\,. \end{array}$$

Eine geometrische Veranschaulichung zeigt: Beide Nebenbedingungen sind im Lösungspunkt aktiv. Damit ergibt sich:

$$\Delta x_1 \;=\; \frac{5}{12}, \quad \Delta x_2 \;=\; -\,\frac{1}{3}, \quad J_{\text{lin}} \;=\; \{1, 2\}.$$

Die Schritte 3.) und 4.) des SQP–Verfahrens liefern schließlich

$$x_1^1 \;=\; \frac{11}{12}, \quad x_2^1 \;=\; \frac{2}{3}, \quad \lambda_1^1 \;=\; \frac{1}{3}, \quad \lambda_2^1 \;=\; \frac{2}{3}\,.$$

Mit diesen Daten ergibt sich nun im zweiten Iterationsschritt eine quadratische Optimierungsaufgabe

$$\begin{array}{ll} \text{Minimiere} & q(\boldsymbol{\Delta x}) \;=\; \Delta x_1^2 \;+\; \dfrac{2}{3}\,\Delta x_2^2 \;-\; \Delta x_1 \;-\; \Delta x_2 \;-\; \dfrac{19}{12} \\[2ex] \text{Nebenbedingungen} & -\dfrac{11}{6}\,\Delta x_1 \;+\; \Delta x_2 \;-\; \dfrac{25}{144} \;\geq\; 0 \\[2ex] & -\dfrac{11}{6}\,\Delta x_1 \;-\; \dfrac{4}{3}\,\Delta x_2 \;-\; \dfrac{41}{144} \;\geq\; 0\,. \end{array}$$

Eine geometrische Darstellung der Nebenbedingungen und der Höhenlinien der Zielfunktion (Ellipsen!) zeigt, daß diesmal im Lösungspunkt nur die zweite Nebenbedingung aktiv ist. Mit dieser Gleichung ($c_2(\Delta x_1, \Delta x_2) = 0$) läßt sich beispielsweise Δx_2 aus der Zielfunktion eliminieren und man erhält eine unrestringierte quadratische Optimierungsaufgabe in einer Variablen (Δx_1). Konkret ergibt sich

$$\Delta x_1 \approx -0.16955, \quad \Delta x_2 \approx 0.019585, \quad J_{\text{lin}} = \{2\}.$$

Das Gleichungssystem zur Berechnung des Lagrange-Parameters λ_2 (Schritt 3 des Algorithmus) ist nun tatsächlich überbestimmt (zwei lineare Gleichungen für eine Unbekannte) - jedoch lösbar! Insgesamt ergibt sich für die neue Näherung

$$x_1^2 = 0.74712, \quad x_2^2 = 0.68625, \quad \lambda_1^2 = 0, \quad \lambda_2^2 = 0.73042\,.$$

Ab diesem Iterationschritt bleibt die Lösungsstruktur in den folgenden Schritten unverändert: Stets ist nur die zweite Nebenbedingung aktiv. Bereits die folgende Iterierte liefert schon Ergebnisse, die mit den exakten Ergebnissen auf zwei bis drei Dezimalstellen übereinstimmen:

$$x_1^3 = 0.70876, \quad x_2^3 = 0.70679, \quad \lambda_1^3 = 0, \quad \lambda_2^3 = 0.70673\,.$$