Moderne Methoden der nichtlinearen endlichdimensionalen Optimierung

Mitschrift[1] zur Vorlesung von Prof. Griewank

Sommersemester 2004

[1]Gefundene Fehler bitte an vigerske@mathematik.hu-berlin.de.

Inhaltsverzeichnis

Allgemeine Voraussetzung

$$\begin{array}{ll} \min & f(x) \\ \text{s.d.} & h(x) = 0 \\ & c(x) \leq 0 \end{array}$$

wobei $f : \mathbb{R}^n \to \mathbb{R}$, $h : \mathbb{R}^n \to \mathbb{R}^m$, $c : \mathbb{R}^n \to \mathbb{R}^p$ und $f, g, h \in C^2(V)$ oder $C^1(V)$ für $V \subseteq \mathbb{R}^n$.

Kapitel 1

Freie Optimierungsprobleme

$m = p = 0$

1.1 Abstiegsminimierung mit Strahlsuche

$$\begin{aligned} x_{k+1} &= x_k + \alpha_k s_k \\ \text{mit } s_k &= -B_k^{-1} g_k \end{aligned}$$

wobei $g_k = \nabla f(x_k)$, $B_k^T = B_k \approx \nabla^2 f(x_k)$.

Abstiegswinkel $\varphi_k := \angle(s_k, -g_k)$ für $B_k \succ 0$, d.h. pos. definit

$$\cos(s_k, -g_k) = \frac{-g_k^T s_k}{\|g_k\| \|s_k\|} \geq \frac{1}{\sqrt{\kappa(B_k)}} \qquad \kappa(B_k) = \|B_k\| \left\|B_k^{-1}\right\|$$

Globale Konvergenz im Sinne von

$$\limsup_{k \to \infty} \|\nabla f(x_k)\| = 0 \qquad \text{oder} \qquad \lim_{k \to \infty} f(x_k) = -\infty$$

folgt aus Kombination von Zoutendijkbedingung

$$D := \sum_{k=0}^{\infty} \cos^2 \varphi_k \geq \sum_{k=0}^{\infty} \frac{1}{\kappa(B_k)}$$

(d.h. $\sup\{\kappa(B_k)\} < \infty$ reicht) mit effektiver Strahlsuche.

Falls zudem

$$\sup \|B_k\| < \infty, \qquad \sup \left\|B_k^{-1}\right\| < \infty$$

folgt sogar

$$\lim_{k \to \infty} x_k = x_*,$$

d.h. Konvergenz zu stationärem Punkt, wo Optimalitätsbedingung 1. Ordnung erfüllt ist.

Falls zudem $\nabla^2 f(x_*) \succ 0$ muß x_* lokaler Minimalpunkt sein und $B_k = \nabla^2 f(x_k)$ führt zu Q-quadratischer Konvergenz, d.h.

$$\|x_{k+1} - x_*\| \leq c \|x_k - x_*\|^2 \qquad \text{für Konstante } c.$$

Falls nur $B_k \approx \nabla^2 f(x_k)$ mittels Sekantenaufdatierung ergibt sich immer noch Q-superlineare Konvergenz, d.h.

$$\lim_{k \to \infty} \frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|} = 0$$

Bemerkung 1.1. Unter geeigneten Voraussetzungen lassen sich quadratische Konvergenz für Newton und superlineare Konvergenz für quasi-Newton auch bei NLP-Lösern, z.B. SQP und IP (Interior Point), beweisen.

1.2 Abstiegsminimierung mit Trust Region-Verfahren

Motivation: Problem bei Strahlsuche: Was tun, wenn $\nabla^2 f(x)$ negative Eigenwerte hat?

1. nur für Iterierte $x = x_n$ (auf Weg)
2. auch am stationären Punkt $x_* = \lim x_n$, d.h. notwendige Optimalitätsbedingung 2. Ordnung ist verletzt.

Antwort 1: Korrektur zu 1.: Ersetze $\nabla^2 f(x_n)$ durch positiv definites B_k, z.B. $B_k = \nabla^2 f(x_n) + \lambda_k I$ mit $\lambda_k > 0$.

Diese Modifikation reicht nicht aus, um 2. zu beseitigen.

Beispiel.

$$\begin{aligned} f(x,y) &= \frac{1}{2}x^2 - \frac{1}{2}y^2 \\ g(x,y) = \nabla f(x,y) &= (x,-y) \\ \nabla^2 f(x,y) &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{ist überall indefinit} \end{aligned}$$

Vom Ausganspunkt (x_0, y_0) mit $x_0 \neq 0$, $y_0 = 0$ ergibt die quasi-Newton Iteration mit

$$B_k = \begin{pmatrix} 1+\lambda_k & 0 \\ 0 & -1+\lambda_k \end{pmatrix}$$

für $\lambda_k \in [2, 100]$ für alle $k \geq 0$

$$\begin{aligned} y_k &= 0 \\ x_{k+1} &= x_k - x_k \frac{1}{1+\lambda_k} = \frac{\lambda_k}{1+\lambda_k} x_k \end{aligned}$$

$\rightsquigarrow \frac{x_{k+1}}{x_k} \in \left[\frac{2}{3}, \frac{100}{101}\right] \rightsquigarrow (x_*, y_*) = \lim_{k\to\infty}(x_k, y_k) = (0,0)$.

Also konvergiert die Folge (x_n, y_n) gegen einen stationären Punkt, der nicht minimal ist.

Antwort zu (1) und (2) ist das *Trust Region Modell*.

Idee: Wähle *Trust Region Radius* $\Delta_k > 0$, so daß

$$f(x_k + s) \approx \Phi_k(s) := f(x) + g(x)^T s + \frac{1}{2} s^T B_k s \qquad \text{für } \|s\|_2 \leq \Delta_k$$

1. *Bestimme* s_k als Lösung von

$$\min \Phi_k(s) \quad \text{s.d. } \|s\| \leq \Delta_k$$

2. *Vergleiche* tatsächliche Funktionsdifferenz (actual reduction)

$$\text{ared}_k = -\left[f(x_k + s_k) - f(x_k)\right] = f(x_k) - f(x_k + s_k)$$

mit vorhergesagter Reduktion (predicted reduction)

$$\text{pred}_k = -\left[\Phi_k(s_k) - f(x_k)\right] = f(x_k) - \Phi_k(s_k) \geq 0$$

durch Berechnung des Quotienten

$$r_k = \frac{\text{ared}_k}{\text{pred}_k}.$$

3. *Akzeptiere* $x_{k+1} = x_k + s_k$, falls $r_k > \eta$, sonst setze $x_{k+1} = x_k$ und $\Delta_{k+1} \leq c_4 \Delta_k$ mit $c_4 \in (0,1)$.

Bemerkung 1.2. Im Unterschied zur Strahlsuche wird bei Rückweisung von $x_k + s_k$ nicht nur die Schrittlänge, sondern auch die Schrittrichtung neu berechnet.

Unter der schwachen Voraussetzung $\sup_k \|B_k\| < \infty$, d.h. B_k's dürfen indefinit und speziell singulär sein, folgt unter gleichen Voraussetzungen wie in 1.1 für Strahlsuche "globale Konvergenz" im Sinne von $\liminf \|g_k\| = 0$. Unter zusätzlichen Voraussetzungen gilt sogar $x_* = \lim_{k\to\infty} x_k$, $g(x_*) = f(x_*) = 0$, $\nabla^2 f(x_*) \succeq 0$.

1.2.1 Lösung des Trust Region Modells (exakt)

Nichtkonstruktive Existenzaussage: Die quadratische Funktion

$$\Phi(s) \equiv f + g^T s + \frac{1}{2} s^T B s$$

hat auf der kompakten Kugel

$$\{s \in \mathbb{R}^n \,|c(s) \leq 0\} \qquad \text{mit } c(s) = \frac{1}{2}\|s\|^2 - \frac{1}{2}\Delta^2$$

mindestens einen globalen Minimalpunkt $s_* \in \operatorname{argmin}\{\Phi(s)\,|c(s) \leq 0\}$.

Fragen:

- Gibt es weitere lokale Minima?
- Wie kann man globales Minimum eindeutig berechnen?

Zur Beantwortung betrachte (theoretische) Variablentransformation

$$s = Qz \qquad g = Qh$$

wobei $B = Q\Lambda Q^T$, $QQ^T = I$, Λ diagonal, so daß

$$\begin{aligned}
\varphi(z) := \Phi(Qz) &= f + g^T Q z + \frac{1}{2} z^T Q^T Q \Lambda Q^T Q z \\
&= f + h^T z + \frac{1}{2} z^T \Lambda z \\
&= f + \sum_{i=1}^{n} \left(h_i z_i + \frac{1}{2} \lambda_i z_i^2 \right) \\
\text{und } c(Qz) &= \frac{1}{2}\|Qz\|^2 - \frac{1}{2}\Delta^2 = c(z)
\end{aligned}$$

Mit anderen Worten: O.B.d.A. kann B als diagonal angenommen werden.

Untersuchung des transformierten Problemes

$$\min \varphi(z) \quad \text{s.d. } c(z) \leq 0$$

KKT-Bedingungen für stationären Punkt $z \neq 0$ ($\rightsquigarrow \nabla c(z) = z \neq 0$, d.h. LICQ erfüllt):

$$\begin{aligned} 0 &= \nabla\varphi(z) + \lambda\nabla c(z) = h + \Lambda z + \lambda z = h + (\Lambda + \lambda I) z \\ 0 &\leq \lambda \\ c(z) &\leq 0 \\ 0 &= \lambda c(z) = \lambda\left(\frac{1}{2}\|z\|^2 - \frac{1}{2}\Delta^2\right) \end{aligned}$$

Fall 1. Es existiert ein lokaler Minimalpunkt z im Inneren, d.h. $c(z) < 0$.

$\rightsquigarrow \lambda = 0$ (wegen Komplementarität)

$$\begin{aligned} \nabla\varphi(z) = h + \Lambda z &= 0 \qquad \text{(notwendige Optimalität 1. Ordnung)} \\ \nabla^2\varphi(z) = \Lambda &\succeq 0 \qquad \text{(notwendige Optimalität 2. Ordnung)} \end{aligned}$$

Da aus der positiven Semidefinitheit von Λ die Konvexität von $\varphi(z)$ folgt, sind diese notwendigen Bedingungen sogar hinreichend für globale Optimalität von z im $\mathbb{R}^n$.

Also gibt es in der Trust Region keine nichtglobalen lokalen Minima.

Letzteres gilt sowohl bezüglich z wie auch für die ursprünglichen Variablen s.

Fall 2. Es existiert ein lokales Minima z am Rand, d.h. $c(z) = 0$.

Notwendige Bedingungen 1. Ordnung:

$$h + (\Lambda + \lambda I) z = 0 \qquad \frac{1}{2}\|z\|^2 - \frac{1}{2}\Delta^2 = 0$$

Notwendige Bedingung 2. Ordnung:

$$\nabla c(z)^T \zeta = z^T\zeta = 0 \quad \Rightarrow \quad \zeta^T\left(\nabla_z^2 L(z,\lambda)\right)\zeta = \zeta^T(\Lambda + \lambda I)\zeta \geq 0$$

wobei $L(z,\lambda) = f + h^T z + \frac{1}{2} z^T \Lambda z + \lambda \frac{1}{2}\left(\|z\|^2 - \Delta^2\right)$

Hinreichend für lokale Optimalität von z ist die strikte Ungleichung

$$\zeta^T(\Lambda + \lambda I)\zeta = \sum_{i=1}^n (\lambda_i + \lambda)\zeta_i^2 > 0$$

für alle $\zeta \neq 0$ mit $z^T\zeta = 0$.

Selbst das reicht nicht für globale Optimalität aus, wie an folgendem Beispiel ersichtlich.

Beispiel. $n = 1$, $\varphi(z) = -\frac{1}{2}\left(z - \frac{1}{2}\right)^2$, $\Delta = 1$

Hinreichende Optimalitätsbedingung ist sowohl für $z = -1$ wie auch $z = 1$ erfüllt, da $z^T\zeta = 0$ nur für $\zeta = 0$ gilt.

Allgemein ($n \geq 1$) existieren maximal zwei lokal minimale Werte. Das globale Minimum erfüllt sogar $\Lambda + \lambda I \succeq 0$ und läßt sich wie folgt charakterisieren.

Zusatzannahme: $h_i \neq 0$ für $i = 1, \ldots, n$

$\rightsquigarrow z = -(\Lambda + \lambda I)^{-1} h$

$\rightsquigarrow \lambda$ ist Nullstelle der Funktion

$$\begin{aligned} \psi(\lambda) = c(z(\lambda)) &= \frac{1}{2}\left\|(\Lambda + \lambda I)^{-1} h\right\|^2 - \frac{1}{2}\Delta^2 \\ &= \frac{1}{2}\sum_{i=1}^{n} \frac{h_i^2}{(\lambda_i + \lambda)^2} - \frac{1}{2}\Delta^2 \end{aligned}$$

Sei $\lambda_1 < \lambda_2 < \ldots < \lambda_n$. $\psi(\lambda)$ hat bis zu $2n$ Nullstellen von denen genau eine $\lambda = \lambda_*$ größer als $-\lambda_1$ sein muß, da $h_1 \neq 0$ vorausgesetzt ist.

Für dieses λ_* ist $(\Lambda + \lambda_* I) = \operatorname{diag}(\lambda_i + \lambda)_{i=1,\ldots,n}$ sogar positiv definit, die hinreichenden Optimalitätsbedingungen 2. Ordnung sind also an $z_* = -(\Lambda + \lambda_* I)^{-1} h$ erfüllt. z_* ist sogar globales Minimum, da für beliebiges ζ mit $c(z_* + \zeta) \leq 0$ gilt

$$\begin{aligned} 0 \geq c(z_* + \zeta) &= \frac{1}{2}\|z_* + \zeta\|^2 - \Delta^2 \\ &= \frac{1}{2}\|z_*\|^2 + z_*^T\zeta + \frac{1}{2}\|\zeta\|^2 - \frac{1}{2}\Delta^2 \\ &= z_*^T\zeta + \frac{1}{2}\|\zeta\|^2 \\ \rightsquigarrow -z_*^T\zeta &\geq \frac{1}{2}\|\zeta\|^2 \\ \rightsquigarrow \varphi(z_* + \zeta) - \varphi(z_*) &= h^T\zeta + \frac{1}{2}(z_* + \zeta)^T \Lambda (z_* + \zeta) - \frac{1}{2} z_*^T \Lambda z_* \\ &= h^T\zeta + z_*^T\Lambda\zeta + \frac{1}{2}\zeta^T\Lambda\zeta \\ &= \zeta^T(h + \Lambda z_*) + \frac{1}{2}\zeta^T\Lambda\zeta \\ &= \zeta^T(-\lambda I) z_* + \frac{1}{2}\zeta^T\Lambda\zeta \\ &\geq \frac{1}{2}\lambda\|\zeta\|^2 + \frac{1}{2}\zeta^T\Lambda\zeta \\ &= \frac{1}{2}\zeta^T(\Lambda + \lambda I)\zeta \\ &> 0 \end{aligned}$$

Satz 1.3. *s_* ist globales Minimum des Problems*

$$\min \Phi(s) = f + g^T s + \frac{1}{2} s^T B s \qquad s.d.\ \|s\| \leq \Delta$$

gdw. für ein $\lambda_* \geq 0$

$$\begin{aligned} Bs_* + g + \lambda_* s_* &= 0, \\ (\Delta - \|s_*\|) \geq 0 &= \lambda_*(\Delta - \|s_*\|) \\ B + \lambda_* I &\succeq 0 \end{aligned}$$

Beweis. Falls $h = Q^T g$ keine verschwindenden Elemente hat, ergibt sich die Aussage aus vorgehenden Betrachtungen für den Fall 1 und 2. Anderfalls betrachte das gestörte Problem mit

$$h_\varepsilon = Q^T g + \varepsilon\mathbf{1} = Q^T(g + \varepsilon Q\mathbf{1}) \quad \text{für } \varepsilon = 0, \mathbf{1} = (1, \ldots, 1)^T$$

Dann benutze Stetigkeitsargument, d.h. für alle hinreichend kleinen positiven $\varepsilon > 0$ hat h_ε keine verschwindenden Elemente, so daß es s_ε und λ_ε gibt mit

$$\|s_\varepsilon\| \leq \Delta \quad \text{und} \quad Bs_* + g + \lambda_\varepsilon s_\varepsilon = 0.$$

Da $\{s_\varepsilon\}$ beschränkt ist, gibt es mindestens einen Häufungspunkt s_*.

Falls es keinen nichtverschwindenden Häufungspunkt gibt, gilt $0 = \lim_{\varepsilon\to 0} s_\varepsilon$, so daß für hinreichend kleine ε alle Minimalpunkte s_ε unbeschränkt sind und deshalb $0 = Bs_\varepsilon + g_\varepsilon$. Dann folgt $\lim_{\varepsilon\to 0} g_\varepsilon = \lim_{\varepsilon\to 0} g + \varepsilon Q\mathbf{1} = 0 \Rightarrow g = 0, B \succeq 0$. Also ist $s_* = 0$ globales Minimum des ungestörten Problemes und der Satz gilt z.B. für $\lambda_* = 0$.

Falls ein nichtverschwindender Häufungspunkt $s_* = \lim_{j\to\infty} s_{\varepsilon_j}$ mit $\lim_{j\to\infty} \varepsilon_j = 0$ existiert, so folgt für die entsprechenden λ_{ε_j}, daß

$$\limsup_{j\to\infty} \left|\lambda_{\varepsilon_j}\right| \le \limsup_{j\to\infty} \frac{\left\|Bs_{\varepsilon_j} + g + \varepsilon_j Q^T \mathbf{1}\right\|}{\left\|s_{\varepsilon_j}\right\|} \le \|B\| + \frac{\|g\|}{\|s_*\|}.$$

Also besitzen auch die λ_{ε_j} einen Häufungspunkt λ_* der mit s_* die obigen Bedingungen erfüllt. $\square$

Falls $B \succ 0$ und deshalb $s^q = -B^{-1}g$ wohl definiert ist, ergibt sich ein glatter Pfad

$$\begin{aligned} s(\lambda) &= -(B+\lambda I)^{-1} g \\ \text{mit } s(0) &= s^q = -B^{-1} g \\ \text{und } s(\lambda) &\approx -\frac{1}{\lambda} g \qquad \text{für } \lambda \gg 0 \end{aligned}$$

Im normalen, nichtdegenerierten Fall löst λ die Skalargleichung

$$\begin{aligned} 0 = \varphi(\lambda) &= \left\|(B+\lambda I)^{-1} g\right\| - \Delta \\ &= \left[g^T (B-\lambda I)^{-2} g\right]^{\frac{1}{2}} - \Delta \\ &\approx \frac{\left|g^T Q e_1\right|}{\lambda + \lambda_1} - \Delta \qquad \text{falls } \lambda + \lambda_1 \gtrsim 0, \text{ d.h. klein positiv} \end{aligned}$$

Für die Nullstellensuche ist die transformierte (annähernd lineare) Funktion

$$\psi(\lambda) = \frac{1}{\varphi(\lambda) + \Delta} - \frac{1}{\Delta} \approx \frac{(\lambda + \lambda_1)}{|g^T Q e_1|} - \frac{1}{\Delta}$$

besser geeignet. Iterative Lösung mit Newton verlangt Ableitungen

$$\varphi'(\lambda) = \frac{d}{d\lambda}\varphi(\lambda) = \frac{\frac{1}{2}(-2)}{\varphi(\lambda)+\Delta} g^T (B+\lambda I)^{-3} g = -\frac{g^T (B+\lambda I)^{-3} g}{\varphi(\lambda)+\Delta} < 0,$$

denn wegen

$$0 = \frac{d}{d\lambda}\left[(B+\lambda I)^{-2}(B+\lambda I)(B+\lambda I)\right] = \frac{d}{d\lambda}\left[(B+\lambda I)^{-2}\right] \cdot (B+\lambda I)^2 + 2(B+\lambda I)^{-2} I (B+\lambda I)$$

ist

$$\frac{d}{d\lambda}\left[(B+\lambda I)^{-2}\right] = -2(B+\lambda I)^{-3}.$$

Entsprechend gilt

$$\psi'(\lambda) = -\frac{\varphi'(\lambda)}{[\varphi(\lambda)+\Delta]^2} = \frac{g^T (B+\lambda I)^{-3} g}{(\varphi(\lambda)+\Delta)^3} > 0,$$

falls $B + \lambda I \succ 0$.

Newton-Iteration (mit geeigneter Stabilisierung)

$$\lambda \leftarrow \lambda - \frac{\psi(\lambda)}{\psi'(\lambda)} = \lambda - \frac{\varphi(\lambda)(\varphi(\lambda)+\Delta)}{\Delta \cdot \varphi'(\lambda)}$$

verlangt an Hauptarbeit die Lösung der linearen Systeme

$$(B+\lambda I)\,s = g \quad \text{und} \quad (B+\lambda I)\,t = s$$

für eine Folge von Multiplikatoren λ. Für eine direkte Lösung bestehen folgende Möglichkeiten:

1. Die einmalige Eigenwertzerlegung $B = Q\Lambda Q^T$ zum Preis von $\approx 4n^3$ Operationen ermöglicht die sukzessive Lösung des transformierten Problemes mit Aufwand $O(n)$ pro Iteration.

2. Eine wiederholte Cholesky-Faktorisierung $B+\lambda I = LDL^T$ mit $D \succ 0$ verlangt jedesmal einen Aufwand von $\approx \frac{1}{6}n^3$ pro Iteration. (Moré Sorenson)

3. Eine einmalige Reduktion von B auf symmetrische Hessenberg-tridiagonale Form $B = QTQ^T$ mit tridiagonaler Matrix T und $Q^TQ = I$ kostet $\frac{2}{3}n^3$ Operationen. (Ähnlichkeitstransformation auf Hessenbergform wird allgemein als erster Schritt bei der Berechnung von Eigenwerten eingesetzt (QR-Algorithmus). Die Berechnung von Q erfolgt durch Householder Matrizen (elementare Reflektoren).)

   In transformierten Systemen gilt dann

$$\begin{aligned}\varphi(\lambda) &= \|\tilde{s}\| - \Delta \quad \text{mit } (T+\lambda I)\,\tilde{s} = h = Q^T g \\ \varphi'(\lambda) &= \frac{\tilde{s}^T\tilde{t}}{\varphi(\lambda)+\Delta} \quad \text{mit } (T+\lambda I)\,\tilde{t} = \tilde{s}\end{aligned}$$

   $\tilde{s}$ und $\tilde{t}$ lassen sich mittels Cholesky-Verfahren mit Aufwand $O(n)$ berechnen.

Bemerkung 1.4. Das Trust-Region-Problem wird heute nur bei Problemen mit mittlerer Größe, d.h. ≈ 300 Variablen, genau gelöst. Bei wirklich großen Problemen kommen folgende Vergröberungen zum Einsatz.

1.2.2 Vereinfachungen einschließlich Steihaug's CG-Variante

Zur Vermeidung jeglicher Gleichungslösung betrachte den *Cauchy-Punkt*

$$s^c = -g \cdot \underset{0\le\alpha\le\Delta/\|g\|}{\operatorname{argmin}} \left(\Phi(-\alpha g)\right) = -g \cdot \underset{0\le\alpha\le\Delta/\|g\|}{\operatorname{argmin}} \left(-\alpha\|g\|^2 + \frac{1}{2}\alpha^2 g^T B g\right)$$

Freier Minimalpunkt an Stelle $\alpha = \frac{\|g\|^2}{g^T B g}$.

Lemma 1.5. *Der Cauchy-Punkt $s = s^c$ verspricht in jedem Fall die Reduktion*

$$\text{pred} = f - \Phi(s) \ge \frac{1}{2}\|g\| \min\left(\Delta, \frac{\|g\|}{\|B\|}\right).$$

Falls $g^T B g \le 0$ folgt unmittelbar $s = -g\frac{\Delta}{\|g\|}$ und

$$f - \Phi(s) = -\left[\underbrace{-\|g\|^2\frac{\Delta}{\|g\|}}_{g^T s} + \frac{1}{2}\underbrace{\frac{g^T B g \Delta^2}{\|g\|^2}}_{s^T B s}\right] \ge \|g\|\,\Delta.$$

Falls $g^T Bg > 0$ und $\alpha = \frac{\|g\|^2}{g^T Bg} \leq \Delta$ gilt, so ist

$$f - \Phi(s) = -\left[-\frac{\|g\|^4}{g^T Bg} + \frac{1}{2} g^T Bg \frac{\|g\|^4}{(g^T Bg)^2}\right] = \frac{1}{2}\frac{\|g\|^4}{g^T Bg} \geq \frac{1}{2}\frac{\|g\|^2}{\|B\|} = \frac{1}{2}\|g\|\left(\frac{\|g\|}{\|B\|}\right).$$

Falls $g^T Bg > 0$ und $\frac{\|g\|^2}{g^T Bg} \geq \frac{\Delta}{\|g\|}$, d.h. $\frac{g^T Bg}{\|g\|^2} \leq \frac{\|g\|}{\Delta}$, gilt, so ist $\alpha = \frac{\Delta}{\|g\|}$, $s = -\frac{g}{\|g\|}\Delta$ und

$$f - \Phi(s) = -\left[-\Delta\|g\| + \frac{1}{2}\frac{\Delta^2}{\|g\|^2} g^T Bg\right] \geq \Delta\|g\| - \frac{1}{2}\frac{\Delta^2\|g\|}{\Delta} = \frac{1}{2}\Delta\|g\|.$$

Idee: Falls der Cauchy s^c noch im Inneren der Trust-Region liegt, entspricht er genau der ersten Iterierten s_1 des Conjugate-Gradient(CG)-Verfahrens, angewandt auf $Bs = -g$, gestartet mit $s_0 = 0$.

Die Iteration kann fortgesetzt werden, bis zum ersten mal

1. eine Richtung ζ_k mit $\zeta_k B\zeta_k \leq 0$ auftritt, oder
2. ein normaler CG-Schritt $s_k - \frac{\nabla\Phi(s_k)\zeta_k}{\zeta_k B\zeta_k}$ die Trust-Region verlassen würde.

In beiden Fällen wird α_k so gewählt, daß $\|s_k + \alpha_k\zeta_k\| = \Delta$ gilt.

Falls weder 1. noch 2. auftreten, wird nach spätestens n Schritten der Minimalpunkt $s_* = -B^{-1}g$ von $\Phi(s)$ erreicht. Letzeres folgt aus folgendem Lemma:

Lemma 1.6. *Das CG-Verfahren mit $s_0 = 0$ verläßt die Trust-Region höchstens genau einmal, da die Norm $\|s_k\| = \|s_k\|_2$ monoton wächst, solange $\zeta_k^T B\zeta_k > 0$ ist.*

Beweis. Die Aussage folgt offenbar aus $s_k^T\zeta_k \geq 0$ für alle $k = 0, 1, \ldots$. Wir zeigen dies durch Induktion über k:

$$s_{k+1} = s_k + \alpha_k\zeta_k$$

Induktionsanfang $k = 0$: $\zeta_0 = -g$, $s_0 = 0 \Rightarrow s_0^T\zeta_0 = 0 \geq 0$

Induktionsschluß:

$$s_{k+1}^T\zeta_{k+1} = s_{k+1}^T\left(-\nabla\Phi(s_{k+1}) + \underbrace{\beta_k}_{\geq 0}\zeta_k\right),$$

wobei s_{k+1} eine Linearkombination der Schritt ζ_j für $j \leq k$ ist. Wegen der Expanding-Subspace-Eigenschaft ist nach dem CG-Theorem der aktuelle Gradient $\nabla\Phi(s_{k+1})$ zu allen Suchrichtungen ζ_j mit $j \leq k$ orthogonal. Also gilt

$$\begin{aligned} s_{k+1}^T\zeta_{k+1} &= s_{k+1}^T\beta_k\zeta_k \\ &= (s_k + \alpha_k\zeta_k)^T\beta_k\zeta_k \\ &= \beta_k\underbrace{s_k^T\zeta_k}_{\geq 0 \text{ nach IV}} + \underbrace{\alpha_k\beta_k\|\zeta_k\|^2}_{\geq 0} \\ &\geq 0. \end{aligned}$$

□

Letzte Vereinfachung: Dog-Leg

Falls $B \succ 0$, so betrachte die stückweise lineare Kurve, die den Cauchy-Punkt s^c und das unbeschränkte Minimum $s = -B^{-1}g$ verbindet. Der Dog-Leg-Schritt s^d ist dann der Punkt auf dieser Kurve mit $\Phi(s^d) \leq \Phi(s^c)$.

1.2.3 Beweis der globalen Konvergenz

Algorithmus 1.7. Betrachte den Algorithmus (nach [NoWr]):

Initialisiere $\Delta_0 > 0$ und $\eta \in [0,1)$ sowie $x_0 \in \mathbb{R}^n$.
for $k = 0, 1, \ldots$ **do**
  Wähle $s_k \in \mathbb{R}^n$, so daß

$$\mathrm{pred}_k = f(x_k) - \Phi_k(s_k) \geq \frac{1}{2} \|g_k\| \cdot \min\left(\Delta_k, \frac{\|g_k\|}{\|B_k\|}\right)$$

  Berechne

$$r_k = \frac{f(x_k) - f(x_k + s_k)}{\mathrm{pred}_k} = \frac{\mathrm{ared}_k}{\mathrm{pred}_k}$$

  if $r_k \geq \eta$ **then**
    setze $x_{k+1} = x_k + s_k$
  else
    $x_{k+1} = x_k$
  end if
  if $r_k < \frac{1}{4}$ **then**
    setze $\Delta_{k+1} = \frac{1}{4}\|s_k\| \leq \frac{1}{4}\Delta_k$
  else if $r_k > \frac{3}{4}$ und $\|s_k\| = \Delta_k$ **then**
    $\Delta_{k+1} = 2\Delta_k$
  else
    $\Delta_{k+1} = \Delta_k$
  end if
end for

Satz 1.8. *Angenommen, die Menge $\{x \in \mathbb{R}^n \,|\, f(x) \leq f(x_0)\}$ ist für gegebenes $x_0 \in \mathbb{R}^n$ beschränkt und f ist auf dieser Niveaumenge Lipschitz-stetig differenzierbar. Falls zudem $\beta = \sup_{k>0} \|B_k\| < \infty$, so hat die erzeugte Iterationsfolge x_k die Eigenschaft*

$$\liminf_{k\to\infty} \|\nabla f(x_k)\| = 0.$$

Insbesondere konvergiert eine Teilfolge zu einem stationären Punkt $x_* \in \mathbb{R}^n$.

Beweis. Entweder haben wir $r_k \geq \frac{1}{4}$ unendlich oft, oder Δ_k wird nur endlich oft nicht um einen Faktor $\frac{1}{4}$ reduziert, so daß $\lim_{k\to\infty} \Delta_k = 0$ gilt.

Im ersten Fall gilt unendlich oft

$$\mathrm{ared}_k \equiv f(x_k) - f(x_{k+1}) \geq \frac{1}{4}\mathrm{pred}_k = \frac{1}{8}\|g_k\| \cdot \min\left(\Delta_k, \frac{\|g_k\|}{\beta}\right).$$

Da die Summe aller akzeptierten $\mathrm{ared}_k \geq 0$ durch $f(x_0) - \min f(x)$ beschränkt ist, muß gelten

$$\liminf_{k\to\infty} \mathrm{ared}_k = 0.$$

Daraus folgt

$$\liminf_{k\to\infty} \|g_k\| \cdot \min\left(\Delta_k, \frac{\|g_k\|}{\beta}\right) = 0$$

und somit entweder

$$\liminf_{k\to\infty} \|g_k\| = 0$$

oder

$$\liminf_{k\to\infty} \Delta_k = 0.$$

Auszuschließen bleibt nur noch der zweite Fall, d.h. wir nehmen o.B.d.A. durch Weglassen von Zwischeniterationen an, daß

$$\lim_{k\to\infty} \Delta_k = 0 \quad \text{und} \quad r_k \le \frac{1}{4} \text{ für alle } k.$$

Für diese (Teil)folge gilt also

$$-g_k^T s_k - \frac{1}{2}L\left\|s_k\right\|^2 \le \text{ared}_k \le \frac{1}{4}\text{pred}_k \le \frac{1}{4}\left(-g_k^T s_k + \underbrace{\frac{1}{2}\beta\left\|s_k\right\|^2}_{\ge -\frac{1}{2}s_k^T B s_k}\right)$$

wobei L die Lipschitzkonstante von $\nabla f(x)$sei. Damit folgt

$$\begin{aligned}
\limsup_{k\to\infty} \frac{3}{4}\frac{\left(-g_k^T s_k\right)}{\Delta_k} &\le \limsup_{k\to\infty} \frac{\left\|s_k\right\|^2}{\Delta_k}\left(\frac{1}{8}\beta + \frac{1}{2}L\right) = 0\\
\Rightarrow \limsup_{k\to\infty} \frac{\text{pred}_k}{\Delta_k} &\le \limsup_{k\to\infty} \frac{-g_k^T s_k + \frac{1}{2}\beta\left\|s_k\right\|^2}{\Delta_k} = 0\\
\Rightarrow \limsup_{k\to\infty} \left\|g_k\right\| \min\left(1, \frac{\left\|g_k\right\|}{\Delta_k\beta}\right) &= 0\\
\Rightarrow \limsup_{k\to\infty} \frac{\left\|g_k\right\|}{\Delta_k} &= 0\\
\Rightarrow \liminf_{k\to\infty} \left\|g_k\right\| &= 0.
\end{aligned}$$

□

Satz 1.9. *Unter den Voraussetzungen von Satz 1.8 erzwingt* $0 < \eta < \frac{1}{4}$ *sogar*

$$\lim_{k\to\infty} g_k = 0.$$

Also ist jeder Häufungspunkt von $\{x_k\}$ *stationär.*

Beweis. Siehe [NoWr]. □

Um auch die notwendigen Optimalitätsbedingungen 2. Ordnung an allen Häufungspunkten zu erzwingen, setze $B_k = \nabla^2 f(x_k)$.

Mit anderen Worten, wir betrachten die Trust-Region-Variante des Newton-Verfahrens. Dann folgt für die exakte Trust-Region-Lösung = Minimalpunkt s_k

$$\begin{aligned}
-f(x_k) + \Phi_k(s_k) &= \min_{s:\|s\|\le\Delta} g_k^T s + \frac{1}{2}s^T\nabla^2 f(x_k)\, s\\
&\le \min_{s:\|s\|\le\Delta} \frac{1}{2}s^T\nabla^2 f(x_k)\, s\\
&\le \min_{s:\|s\|\le\Delta} \frac{1}{2}\|s\|^2\lambda_{\min}\\
&\le \frac{1}{2}\Delta^2\lambda_{\min},
\end{aligned}$$

wobei $\lambda_{\min} = \lambda_{\min}\left(\nabla^2 f(x_k)\right)$ den kleinsten Eigenwert von $\nabla^2 f(x_k)$ bezeichne.

Also gilt

$$\text{pred}_k \ge -\frac{1}{2}\Delta^2\lambda_{\min} > 0 \quad\Leftrightarrow\quad \lambda_{\min} < 0.$$

Satz 1.10. *Unter den Voraussetzungen von Satz 1.9 impliziert*

$$B_k = \nabla^2 f(x_k) \quad \textit{und} \quad s_k = \operatorname*{argmin}_{s:\|s\|\leq\Delta} \Phi_k(s),$$

daß

$$\nabla^2 f(x_*) \succeq 0$$

für alle Häufungspunkte x_* der Folge $\{x_k\}$.

Beweis. Siehe [JaSt]. □

Bemerkung 1.11. Da $\|s\| = \|s\|_2$ sehr stark von der Skalierung der Variablen $e_i^T x$, d.h. der Einheiten in denen diese gemessen werden, abhängt, benutzt man oft Trust Regions der Form

$$\|s\| = \|Ds\|_2,$$

wobei $D = \operatorname{diag}(d_j)_{j=1}^n$ mit $d_j > 0$ oder

$$\begin{aligned} \|s\| &= \|Ds\|_1 = \sum_{i=1}^n |d_i s_i| \\ \text{oder } \|s\| &= \|Ds\|_\infty = \max_{1\leq i\leq n} |d_i s_i|. \end{aligned}$$

Diese ergeben ein QP der Form $\min \Phi_k(s)$ s.d. $\|s\| \leq \Delta$, siehe Abschnitt 2.1.2.

Allgemein ist es schwierig, eine geeignete Matrix D zu finden.

Kapitel 2

Quadratische Programmierung

2.1 Definition, Motivation, Beispiele

$$
\text{(QP)} \qquad \begin{array}{llll}
\min & f(x) & \equiv g^T x + \frac{1}{2} x^T B x & \text{(quadr. Zielfunktion)} \\
\text{s.d.} & a_i^T x = b_i & \text{für } i \in \mathcal{G} & \text{(lineare Gleichheitsrestriktionen)} \\
 & a_i^T x \leq b_i & \text{für } i \in \mathcal{U} & \text{(lineare Ungleichungen)}
\end{array}
$$

Ein (QP) heißt konvex, falls $B \succeq 0$ und streng konvex, falls sogar $B \succ 0$.

Frage: Warum nicht $a_i^T x \leq b_i$ und $-a_i^T x \leq -b_i$ statt $a_i^T x \leq b_i$?

Antwort: Da beide Ungleichungen kolineare Gradienten a_i und $-a_i$ haben, geht auf jeden Fall LICQ (Linear Independent Constraint Qualification) verloren. $\Rightarrow$ Unbestimmte Multiplikatoren

Frage: Warum nicht $A_{\mathcal{G}} x = b_{\mathcal{G}}$ mit $A_{\mathcal{G}} = \left(a_i^T\right)_{i \in \mathcal{G}}$, $b_{\mathcal{G}} = (b_i)_{i \in \mathcal{G}}$ benutzen, um mit dem impliziten Funktionentheorem einen Teil der Variablen x_i zu eliminieren?

Antwort: Zerstört typischerweise Dünnbesetztheit, falls vorhanden!

2.1.1 Orthogonale Projektion

$B = I$, $g = -x_*$. Dann ist

$$
f(x) = -x_*^T x + \frac{1}{2} \|x\|^2 = \frac{1}{2} \|x - x_*\|^2 - \underbrace{\frac{1}{2} \|x_*\|^2}_{\text{konstant}}
$$

$$
M \equiv \left\{ x \in \mathbb{R}^n \,\middle|\, a_i^T x = b_i \text{ für } i \in \mathcal{G},\ a_i^T x \leq b_i \text{ für } i \in \mathcal{U} \right\}
$$

Die Lösung ist nur unwesentlich einfacher als allgemeines konvexes QP. Anwendung von projezierten Gradienten auf konvexe QP macht also wenig Sinn.

2.1.2 Trust Regions bezüglich l_∞ oder l_1 Norm.

g und B beliebig und für $i = 1, \ldots, n$

l_∞-Norm ($2n$ Ungleichungen):

$$
\begin{array}{rcl}
e_i^T x & \leq & b_i + \Delta \\
-e_i^T x & \leq & -b_i + \Delta
\end{array}
\quad \Leftrightarrow \quad |x_i - b_i| \leq \Delta
$$

l_1-Norm (2^n-Ungleichungen):

$$\sum_{i=1}^{n} d_i \left(e_i^T x - b_i\right) \leq \Delta$$

für beliebiges $d = (d_1, \ldots, d_n) \in \{-1, 1\}^n$

Warnung: Sphärisches Trust Region Problem bzgl. $\|x\| = \|x\|_2$

$$\min f(x) \quad \text{s.d.} \quad \|x\|_2^2 \equiv \sum_{i=1}^{n} x_i^2 \leq \Delta$$

ist kein QP, da dieses keine quadratischen Nebenbedingungen zulässt.

2.1.3 QP als lokales Modell für $\min f(x)$, s.d. $c(x) = 0$

In Newtonähnlicher Form bestimme Schritt s am Punkt x als Lösung des QP

$$\min \quad \underbrace{\nabla f(x)^T s + \frac{1}{2} s^T B s}_{\text{quadratisches Zielfunktional}} \qquad \text{s.d.} \qquad \underbrace{c(x) + \nabla c(x) s = 0}_{\text{linearisierte Nebenbedingungen}} \quad .$$

Frage: Können wir $B = \nabla^2 f(x)$ setzen, d.h. das Zielfunktional durch Taylorreihe zweiter Ordnung von $f(x)$ selbst ersetzen?

Antwort 0: Falls $m = \dim(c) = n$, d.h. wir haben $m = n$ aktive Beschränkungen, spielt B keine Rolle, da lediglich Zulässigkeit durch $c(x) = 0$ herzustellen ist. (Ja!) Linearisierte Nebenbedingungen ergeben quadratische Konvergenz unter den üblichen Voraussetzungen, d.h. $\nabla c(x) = c'(x)$ ist Lipschitz und regulär.

Antwort 1: Wenn $m = 0$, das Problem also unbeschränkt, ergibt $B \equiv \nabla^2 f(x)$ Newton's Methode zur Lösung von $\nabla f(x) = 0$ und damit wiederum quadratische Konvergenz. (Ja!)

Antwort 2: Nein! Aus folgendem Beispiel ergibt sich, daß $B = \nabla^2 f(x)$ im allgemeinen Fall nur lineare Konvergenz erzielen kann (Mit anderen Worten: Die 2. Ableitungen (Krümmungen) der Restriktionen können nicht vernachlässigt werden, sondern müssen in B eingehen.).

Beispiel.

$$\begin{aligned} \min \quad & \frac{1}{2}(x_1 - 1)^2 + \frac{1}{2}(x_2 - 1)^2 \\ \text{s.d.} \quad & \sqrt{x_1^2 + x_2^2} - 1 = 0 \end{aligned}$$

Lokales QP mit $B = \nabla^2 f(x)$:

$$\begin{aligned} \min \quad & \frac{1}{2}(x_1 + s_1 - 1)^2 + \frac{1}{2}(x_2 + s_2 - 1)^2 \\ \text{s.d.} \quad & \sqrt{x_1^2 + x_2^2} - 1 + \frac{x_1 s_1 + x_2 s_2}{\sqrt{x_1^2 + x_2^2}} = 0 \end{aligned}$$

Linearisierte Restriktionen entsprechen einer Tangente am Kreis.

Leicht (elementar, aber etwas mühsam) zu zeigen: Die resulierenden Iterierten $x^{(k+1)} = x^{(k)} + s^{(k)}$ konvergieren nur linearen gegen $\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$.

Bemerkung 2.1. Auch lineare Konvergenz ist keinesfalls sicher, die linearisierten Nebenbedingungen können sogar inkonsistent sein, so daß das lokale QP eine leere zulässige Menge hat. Es muß daher *relaxiert* werden. Diese Beobachtung betrifft nur die Restriktionen und gilt deshalb für beliebiges B.

"Herleitung" des "richtigen" B

Wir wollen schnelle, möglichst quadratische, Konvergenz zur Lösung der KKT Gleichung

$$\nabla f(x) + \sum_{i=1}^{m} \lambda_i \nabla c_i(x) = 0, \qquad c_i(x) = 0 \text{ für } i = 1, \ldots, m,$$

immer unter der Voraussetzung von Gleichheitsnebenbedingungen.

Nach einem Satz von Dennis & Moré konvergiert eine Folge (x_k, λ_k) genau dann superlinear gegen eine Lösung (x_*, λ_*), wenn die Korrekturen $(\Delta x_k, \Delta\lambda_k) = (x_k - x_{k-1}, \lambda_k - \lambda_{k-1})$ asymptotisch gegen die Newtonschritte (s_k, σ_k) konvergieren, d.h.

$$\lim_{k\to\infty} \frac{\|\Delta x_k - s_k\| + \|\Delta\lambda_k - \sigma_k\|}{\|\Delta x_k\| + \|\Delta\lambda_k\|} = 0.$$

Die (s_k, σ_k) sind Lösung von

$$\begin{bmatrix} \nabla_x^2 L(x_k, \lambda_k) & \nabla c(x_k)^T \\ \nabla c(x_k) & 0 \end{bmatrix} \begin{bmatrix} s_k \\ \sigma_k \end{bmatrix} = \begin{bmatrix} -\nabla f(x_k) - \sum_{i=1}^m \lambda_i \nabla c_i(x_k) \\ -c(x_k) \end{bmatrix} = \begin{bmatrix} -\nabla_x L(x_k, \lambda_k) \\ -c(x_k) \end{bmatrix}$$

$(L(x, \lambda) = f(x) + \sum_{i=1}^m \lambda_i c_i(x))$.

Für $x = x_k$, $s = s_k$, $g = \nabla f(x_k)$ und $\hat{\lambda} = \lambda_k + \sigma_k$ ergibt sich äquivalenterweise mit $B = \nabla_x^2 L(x, \lambda)$, $A = \nabla c(x)$

$$\begin{bmatrix} B & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} s \\ \hat{\lambda} \end{bmatrix} = \begin{bmatrix} -g \\ -c(x) \end{bmatrix},$$

wegen $A^T\sigma + A^T\lambda = A^T\hat{\lambda}$, d.h. $\sum_{i=1}^n \lambda_i \nabla c_i$ kann von rechts nach links gebracht werden und $\hat{\lambda}$ ersetzt σ.

Falls $B \succ 0$, so ist dies äquivalent zu der Aussage, daß s Minimalpunkt des folgenden QP ist

$$\begin{aligned} \min \quad & g^T s + \frac{1}{2} s^T B s \\ \text{s.d.} \quad & c + As = 0 \end{aligned}$$

D.h. $B = \nabla_x^2 L(x, \lambda)$ für "geeignetes" $\lambda \in \mathbb{R}^m$ ist richtiges B. D.h. die Krümmungen $\nabla^2 c$ der Nebenbedingungen werden gewichtet in den quadratischen Term des lokalen QP Modelles eingebracht.

Bemerkung 2.2. Bei SLP (Sukzessives LP) wird das Zielfunktional entweder überhaupt nicht verändert, oder linearisiert, Konvergenz ist deshalb fast immer nur linear.

Bemerkung 2.3. Auch bei NLPs mit Ungleichungen schien lokale Modellierung mit linearisierten Restriktionen und $B = \nabla_x^2 L$ bislang optimal. (Jetzt wird vermutet, daß die volle Lösung des lokalen QP mit Ungleichungen häufig zu aufwendig ist.)

2.2 Bezeichnungen und Schrittberechnung

Notation. *Aktive Menge*

$$A(x) = \mathcal{G} \cup \left\{ i \in \mathcal{U} \,\middle|\, a_i^T x = b_i \right\}$$

Notation. Die *Arbeitsmenge*

$$W \subseteq A(x) \quad \text{mit } \mathcal{G} \subseteq W$$

ergibt sich nicht eindeutig aus x, sondern dem Iterationsverlauf. $m := |W|$.

Notation. $A_W \in \mathbb{R}^{m\times n}$, $b_W \in \mathbb{R}^m$,

$$\begin{array}{rcl} A_W & := & \{a_i\}_{i\in W} \\ b_W & := & (b_i)_{i\in W} \end{array}$$

Wegen $W \subseteq A(x)$ gilt $A_W x - b_W = 0$.

Notation. $Z_W \in \mathbb{R}^{(n-m)\times n}$, so daß

$$\det\left(\begin{bmatrix} A_W \\ Z_W \end{bmatrix}\right) \neq 0, \quad A_W Z_W^T = 0,\ Z_W A_W^T = 0.$$

Z_W existiert, vorausgesetzt alle auftretenden A_W haben vollen Rang m.

Notation. $Y_W \in \mathbb{R}^{m\times n}$, so daß

$$\det\left(\begin{bmatrix} Y_W \\ Z_W \end{bmatrix}\right) \neq 0$$

$\Rightarrow \det\left(A_W Y_W^T\right) \neq 0$.

2.2.1 Nullraumverfahren zur Berechnung von s_W

Der Vektor

$$s_W = \operatorname{argmin}\left\{(g+Bx)\,s + \frac{1}{2}s^T Bs \,\middle|\, A_W s = \hat{b}\right\}$$

löst nach KKT die Gleichung

$$\begin{bmatrix} B & A_W^T \\ A_W & 0 \end{bmatrix}\begin{bmatrix} s \\ \hat{\lambda} \end{bmatrix} = \begin{bmatrix} -(g+Bx) \\ \hat{b} \end{bmatrix} \tag{2.1}$$

Sei $\hat{g} = g + Bx$.

Die Zerlegung

$$s_W = Y_W^T s_Y + Z_W^T s_Z$$

ergibt durch Einsetzen in die zweite Gleichung von (2.1) folgendes Gleichungssystem zur Bestimmung von s_Y:

$$A_W Y_W^T s_Y + \underbrace{A_W Z_W^T}_{=0} s_Z = \hat{b}. \tag{2.2}$$

Einsetzen in die erste Gleichung von (2.1) ergibt

$$BY_W^T s_Y + BZ_W^T s_Z + A_W^T \hat{\lambda}_W = -\hat{g}.$$

Multiplikation von links mit Z_W ergibt ein Gleichungssystem zur Bestimmung von s_Z,

$$Z_W BZ_W^T s_Z = -Z_W\hat{g} - Z_W BY_W^T s_Y - \underbrace{Z_W A_W^T}_{=0}\hat{\lambda}_W, \tag{2.3}$$

während man durch Multiplikation von links mit Y_W ein Gleichungssystem zur Bestimmung von $\hat{\lambda}_W$ erhält:

$$Y_W A_W^T \hat{\lambda}_W = -Y_W\hat{g} - Y_W BY_W^T s_Y - Y_W BZ_W^T s_Z. \tag{2.4}$$

Gelöst werden müssen also die linearen Systeme (2.2), (2.3) und (2.4) mit den Matrizen $A_W Y_W^T \in \mathbb{R}^{m\times m}$, $Z_W BZ_W^T \in \mathbb{R}^{(n-m)\times(n-m)}$ und $Y_W A_W^T \in \mathbb{R}^{m\times m}$.

Dazu bieten sich QR und eine LV basierte Faktorisierungvariante an.

QR

$$\begin{aligned} A_W &= \underbrace{\begin{bmatrix} * & 0 \end{bmatrix}}_{U_W} \begin{bmatrix} Y_W \\ Z_W \end{bmatrix} & U_W \text{ eine untere Dreiecksmatrix} \\ \Leftrightarrow A_W^T &= \underbrace{\begin{bmatrix} Y_W^T & Z_W^T \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} * \\ 0 \end{bmatrix}}_{R} & R \text{ eine obere Dreiecksmatrix} \end{aligned}$$

$Q^TQ = QQ^T = I \Rightarrow Y_W Z_W^T = 0,\ Y_W Y_W^T = I_m,\ ZZ^T = I_{n-m}$
Für die Matrizen in den linearen Systemen (2.2), (2.3) und (2.4) ist dann

- $A_W Y_W^T = U_W$ unterhalbs dreiecksformig und $\det(A_W) = \det(U_W)$,
- $Y_W A_W^T = U_W^T$ oberhalb dreiecksformig, und
- $Z_W B Z_W^T$ "normalerweise" positiv definit und hat deshalb eine Cholesky-Faktorisierung $Z_W^T B Z_W = LDL^T$ mit $D \succ 0$ diagonal.

Vorteil: Numerische Stabilität ohne jegliche Pivotierung.

Nachteil: Der Speicherbedarf für U_W, Y_W, Z_W und B ist $\frac{1}{2}m^2 + n^2 + \frac{1}{2}n^2 = \frac{1}{2}m^2 + \frac{3}{2}n^2$, während der Speichbedarf von A und B höchstens $mn + \frac{1}{2}n^2$ beträgt, wobei sich die Dünnbesetztheit in A_W nicht auf die Faktoren U_W, Y_W, Z_W und schon garnicht auf die Projektionen $Z_W B Z_W^T$ usw. überträgt.

LV

$$P_W A P_W' = L_W \begin{bmatrix} U_W & N_W \end{bmatrix}$$

mit unterer Dreiecksmatrix L_W, oberer Dreiecksmatrix U_W, Zeilenpermutationen P_W und Spaltenpermutationen P_W'.

$$\Rightarrow \begin{bmatrix} Y_W \\ Z_W \end{bmatrix} = \begin{bmatrix} I_m & 0 \\ -N_W^T U_W^{-T} & I_{n-m} \end{bmatrix}.$$

Probe:

$$\begin{aligned} A_W Z_W^T &= L_W \begin{bmatrix} U_W & N_W \end{bmatrix} \begin{bmatrix} -U_W^{-1} N_W \\ I \end{bmatrix} = L_W \cdot 0 = 0 \\ A_W Y_W^T &= L_W U_W \text{ ist regulär} \end{aligned}$$

Ausdruck für $Z_W B Z_W^T$ ist

$$\begin{bmatrix} -N_W^T U_W^{-T} & I \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{12}^T & B_{22} \end{bmatrix} \begin{bmatrix} -U_W^{-1} N_W \\ I \end{bmatrix} = B_{22} - N_W^T U_W^{-T} B_{12} - B_{12}^T U_W^{-1} N_W + N_W^T U_W^{-T} B_{11} U_W^{-1} N_W$$

Vorteil: $Y_W = I$ ist trivial und Z_W braucht nicht explizit berechnet zu werden. Der Speicherbedarf von U_W, L_W, N_W und B beträgt $4m + \frac{1}{2}n^2$ und ist im vollbesetzen Fall minimal. Im dünnbesetzten Fall läßt sich durch geeignete Pivotisierung der Speicherplatzbedarf eventuell drastisch reduzieren. Das gilt dann auch für die Anzahl der arithmetischen Operationen.

Nachteil: Wenn wegen Änderungen in W Zeilen von A_W herausfallen oder hinzukommen, muß schon aus Gründen der Stabilität eventuell aufwendig nachpivotisiert werden (siehe [KiSc]).

Bemerkung 2.4. Vorteil beider Optionen (QR und LV) ist, daß Hinzunahme oder Entfernung eines Elementes in W mit Aufwand der Ordnung n^2 auf Ebene der Faktorisierungen nachvollzogen werden können. Der Gesamtaufwand für die QP-Lösung ist deshalb $O\left(n^2 \cdot \text{Schrittzahl}\right)$. Die Schrittzahl kann theoretisch exponentiell wachsen (z.B. wie 2^n, siehe Klee-Winty). Folgendes Verfahren ist eine Verallgemeinerung des Simplex-Verfahrens.

2.3 Aktive Mengen Methode für konvexes QP

Algorithmus 2.5 (Skizze). Ausgehend von $x_0 \in M$ mit $W_0 \subset A(x_0)$ erzeuge eine Folge

$$x_{k+1} = x_k + \alpha_k s_k$$

wobei

$$s_k = s_{W_k}(x_k)$$

und jeweils gilt $\alpha_k < 1$, falls der volle Schritt aus M führt. Dann wird $m_{k+1} = |W_{k+1}| > |W_k| = m_k$, da mindestens eine Restriktion aktiv wird.

$\alpha_k = 1$, falls der Schritt in M bleibt. Dann werden die Multiplikatoren $\hat{\lambda}_k$ geprüft.

Gilt $e_i^T \hat{\lambda}_k \geq 0$ für alle $i \in W_k \cap \mathcal{U}$, so ist ein Optimum erreicht. Andernfalls wird ein solches i aus W entfernt.

Zu zeigen:

1. Die Berechnung von α_k (einfach).
2. s_k ist Abstiegsrichtung, falls $s_k^T A_{W_k} = b_k \Leftrightarrow \hat{b}_k = 0$.
3. $W_{k+1} = W_k \setminus \{i\}$ ergibt eine Abstiegsrichtung, wenn $e_i^T \hat{\lambda}_k < 0$.

1.
$$\alpha_k = \min_{\substack{i \in U \setminus W_k \\ a_i^T s_k > 0}} \left(1, \frac{b_i - a_i^T x}{a_i^T s_k}\right)$$

2.
$$\begin{bmatrix} B & A_{W_k}^T \\ A_{W_k} & 0 \end{bmatrix} \begin{bmatrix} s_k \\ \hat{\lambda}_k \end{bmatrix} = - \begin{bmatrix} \hat{g}_k \\ 0 \end{bmatrix},$$

$\hat{g}_k = g + Bx_k$ der aktuelle Gradient.
Multiplikation von links mit $\begin{bmatrix} s_k^T & 0 \end{bmatrix} \in \mathbb{R}^{n+n}$ ergibt

$$s_k^T B s_k + \underbrace{s_k^T A_{W_k}}_{=0} \hat{\lambda}_k = -\hat{g}_k^T s_k,$$

d.h. $\hat{g}_k^T s_k = -s_k^T B s_k$.
Unter der Voraussetzung $B \succ 0$ oder zumindest $z_k B z_k^T \succ 0$ für alle vorkommenden W_k ist $\hat{g}_k^T s_k < 0$.

3. $x = x_{k+1}$ ist Lösung des Problems

$$\min \left\{ g^T x + \frac{1}{2} x^T B x \,|A_W x = b_W \right\}$$

und erfüllt deshalb die KKT-Bedingung

$$g + Bx + A_W^T \hat{\lambda}_W = 0$$

mit $\text{Rang}(A_W) = m = |W_k|$.
Falls ein $\hat{\lambda}_i = e_i^T \hat{\lambda}_W < 0$ ist, setze $\hat{W} = W \setminus \{i\}$ und finde eine Lösung s mit $A_{\hat{W}} s = 0$, aber $a_i^T s = -1$. Dann gilt

$$(g + Bx)^T s = -\hat{\lambda}_W^T A_W s = -\left(-\hat{\lambda}_i\right) = \hat{\lambda}_i < 0,$$

d.h. eine Lösung des durch $\hat{W}$ definierten Problemes führt zu einer weiteren Reduktion.

Es gibt zwei Arten von Schritten, je nachdem ob eine bindende (blockierende) Restriktion erreicht wird oder nicht:

$$x_k + s_k \in \operatorname{argmin}\left\{ f(x) \left| a_i^T x = b_i, i \in W_k \right. \right\}.$$

Falls $\{j\}$ bindend ist, so wird $W_{k+1} = W_k \cup \{j\}$ und ein neuer Schritt berechnet.

Diese Erweiterung von W_k kann höchstens einmal in Folge passieren, da die Gradienten $\{a_i\}_{i\in W_k} \subset \{a_i\}_{i\in A(x_k)}$ nach Voraussetzung linear unabhängig sind. Wird ein $x_{k+1} = x_k + s_k$ mit $\alpha_k = 1$ erreicht, so ist x_{k+1} entweder <u>die</u> Lösung des konvexen QP (da $\det B \neq 0$) oder es gilt:

$$\begin{aligned} \nabla f(x) + \lambda^T \nabla c(x) &= 0 \\ g(x) + B^T x + \lambda^T A_W^T &= 0 \qquad (\text{mit } \lambda_i \geq 0 \text{ falls } g_i(x) \leq 0), \end{aligned}$$

aber mindestens ein Lagrange-Multiplikator λ_i mit $i \in \mathcal{U}$ ist negativ. Dann wird $W_{k+1} = W_k \setminus \{i\}$ gesetzt und der nächste Schritt berechnet.

Problem: Zumindest theoretisch ist es möglich, daß wiederholt $\alpha_k = 0$ ist und somit keine Reduktion in f erzielt wird. Dabei können Restriktionen wiederholt in die aktive Menge W_k eingeführt und wieder rausgenommen werden. Dieses "cycling" Phänomen kann schon in LP auftreten. Es wird meist von Algorithmen ignoriert und durch Rundungsfehler unterbrochen.

2.4 QP Algorithmus nach W&N

Algorithmus 2.6.

Bestimme zulässigen Anfangspunkt x_0, setze $W_0 \subseteq A(x_0)$ mit $\mathcal{G} \subseteq W_0$.
for $k = 0, 1, \ldots$ **do**
  Berechne $s_k = s_{W_k}$ und $\hat\lambda_k = \hat\lambda_{W_k}$

$$\begin{bmatrix} B & A_{W_k}^T \\ A_{W_k} & 0 \end{bmatrix} \begin{bmatrix} s_{W_k} \\ \hat\lambda_{W_k} \end{bmatrix} = -\begin{bmatrix} g + Bx_k \\ (b - Ax_k)_{W_k} \end{bmatrix}.$$

  if $s_k = 0$ **then**
    Betrachte $\hat\lambda_k$.
    if $e_i^T \hat\lambda_k \geq 0$ für alle $i \in W_k \cap \mathcal{U}$ **then**
      Stop mit $x_* = x_k$
    else
      $i_* \equiv \operatorname{argmin}\left\{ e_i^T \hat\lambda_k \left| i \in W_k \cap \mathcal{U} \right. \right\}$, $W_{k+1} = W_k \setminus \{i_*\}$, $x_{k+1} = x_k$
    end if
  else
    Berechne α_k wie oben spezifiziert
    Setze $x_{k+1} = x_k + \alpha_k s_k$
    if $\alpha_k < 1$ **then**
      Setze $W_{k+1} = W_k \cup \{i\}$ für bindende Restriktion $i \in \mathcal{U} \setminus W_k$ mit $a_i^T x_{k+1} = b_i$
    end if
  end if
end for

Satz 2.7. *Vorausgesetzt, alle $A_k \equiv A_{W_k}$ haben vollen Rank und alle $\alpha_k > 0$ für $k = 1, 2, \ldots$.*

Dann konvergiert Algorithmus 2.6 in endlich vielen Schritten zum Minimalpunkt des streng konvexen QP

$$\begin{array}{lll} \min & f(x) \equiv g^T x + \frac{1}{2} x^T B x & (B \succ 0) \\ s.d. & a_i^T x = b_i & \text{für } i \in \mathcal{G} \\ & a_i^T x \leq b_i & \text{für } i \in \mathcal{U} \end{array}$$

Beweis. Wegen vorausgesetzten $\alpha_k > 0$ bilden die $f(x_k)$ für $k = 1, \ldots, n$ eine streng monoton fallende Folge.

Mindestens jede n-te Iterierte ist eindeutiger Minimalpunkt eines Problemes $\min f(x)$ s.d. $a_i^T x = b_i$ für $i \in W$.

Es gibt aber maximal $2^{|\mathcal{U}|}$ unterschiedliche Mengen W. Also ist die Gesamtzahl der Schritte durch $n \cdot 2^{|\mathcal{U}|}$ beschränkt und damit endlich. □

Bemerkung 2.8. Wegen des Klee&Minty Beispiels kann die Schrittzahl des Simplex-Algorithmus und damit auch des Algorithmus 2.6 tatsächlich exponentiell mit der Zahl der Variablen bzw. Beschränkungen wachen. In der Praxis erwartet man eher $O(n|\mathcal{U}|)$ Schritte.

Frage: Wie läßt sich ein zulässiger Anfangspunkt $x_0 \in \mathbb{R}$ bestimmen?

Antwort: Phase I Prozedur, d.h. Lösung eines Hilfs-LP oder -QP

$$\begin{array}{ll}
\min & f(z) = \sum_{i=1}^{m} z_i^q \qquad m = |\mathcal{G}| + |\mathcal{U}|, \; q \in \{1, 2\} \\
\text{s.d.} & z_i \geq 0 \qquad i = 1, \ldots, m \\
& a_i^T x_i + z_i \operatorname{sign}(b_i) = b_i \qquad i \in \mathcal{G} \\
& a_i^T x_i - z_i \leq b_i \qquad i \in \mathcal{U}.
\end{array}$$

Bemerkung 2.9. Ein zulässiger Punkt für dieses relaxierte Problem ist $x = 0$, $z_i = |b_i|$ für $i \in \mathcal{G}$, $z_i = \max\{0, -b_i\}$ für $i \in \mathcal{U}$.

Die Iterierte (x_k, z_k) liefert für das ursprüngliche QP einen zulässigen Punkt, gdw. $f(z_k) = f(0) = 0$ erreicht wird.

Kapitel 3

Sequentielle Quadratische Programmierung (SQP)

$$\min f(x) \qquad \text{s.d.} \quad c(x) = 0 \tag{3.1}$$

mit $f, g \in C^{2,1}$, d.h. 2-mal stetig differenzierbar mit Lipschitz-stetiger 2. Ableitung, $x \in \mathbb{R}^n$.

Idee: Modelliere das Problem am aktuellen Iterationspunkt mittels eines QP-Subproblems und benutze dessen Lösung als nächsten Iterationspunkt.

Herausforderung: Die Wahl des QP-Subproblemes.

3.1 Basis Modell von SQP

Bezeichnungen:

$$\begin{aligned} L(x,\lambda) &= f(x) + \lambda^T c(x) \\ B(x,\lambda) &= \nabla^2_{xx} L(x,\lambda) \\ A(x)^T &= (\nabla c_1(x), \nabla c_2(x), \ldots, \nabla c_m(x)) \end{aligned}$$

Für einen Iterationspunkt (x_k, λ_k),

$$\begin{aligned} B_k &:= B(x_k, \lambda_k) \\ A_k &:= A(x_k, \lambda_k). \end{aligned}$$

Sei p_k eine Lösung von

$$\begin{aligned} \min_p \quad & \nabla f(x_k) p + \frac{1}{2} p^T B_k p \\ \text{s.d.} \quad & A_k p + c(x_k) = 0 \end{aligned} \tag{3.2}$$

für den aktuellen Iterationspunkt x_k. Dann

$$\begin{aligned} x_{k+1} &= x_k + p_k \\ \lambda_{k+1} &= ? \end{aligned}$$

Durch die Wahl von B_k entstehen verschiedene SQP-Verfahren.

Voraussetzungen:

1. LICQ, d.h. $\nabla c_i(x_k)$, $i = 1, \ldots, m$, sind linear unabhängig. Mit anderen Worten: A_k hat vollen Zeilenrank.
2. $d^T B_k d > 0$ für alle $d \neq 0$ mit $A_k^T d = 0$.

Das Problem (3.2) hat eine eindeutige Lösung (p_k, μ_k), welches die KKT-Bedingungen erfüllt:

$$\begin{bmatrix} B_k & A_k^T \\ A_k & 0 \end{bmatrix} \begin{bmatrix} p_k \\ \mu_k \end{bmatrix} = \begin{bmatrix} -\nabla f(x_k) \\ -c(x_k) \end{bmatrix}. \tag{3.3}$$

Andererseits: Das KKT von (3.1) ist

$$\begin{aligned} F(x,\lambda) := \nabla f(x) + A^T(x)\lambda &= 0 \\ c(x) &= 0 \end{aligned}$$

Wenden wir auf $F(x,\lambda) = 0$ das gewöhnliche Newton-Verfahren im Punkt (x_k, λ_k) an, so erhalten wir

$$F(x_k,\lambda_k) + \frac{\partial F(x_k,\lambda_k)}{\partial(x,\lambda)} \begin{pmatrix} \Delta x_k \\ \Delta \lambda_k \end{pmatrix} = 0,$$

d.h.

$$\begin{aligned} \begin{bmatrix} B_k & A_k^T \\ A_k & 0 \end{bmatrix} \begin{bmatrix} \Delta x_k \\ \Delta \lambda_k \end{bmatrix} &= \begin{bmatrix} -\nabla f(x_k) - A_k^T \lambda_k \\ -c(x_k) \end{bmatrix} \\ \text{bzw. } \begin{bmatrix} B_k & A_k^T \\ A_k & 0 \end{bmatrix} \begin{bmatrix} \Delta x_k \\ \lambda_k + \Delta \lambda_k \end{bmatrix} &= \begin{bmatrix} -\nabla f(x_k) \\ -c(x_k) \end{bmatrix} \end{aligned} \tag{3.4}$$

Wenn $p_k = \Delta x_k$ und $\mu_k = \lambda_{k+1} = \lambda_k + \Delta \lambda_k$, so ist (3.3) äquivalent zu (3.4).

Zeigen: (3.2) ist äquivalent zu

$$\begin{aligned} \min_p \quad & \nabla_x L(x_k,\lambda_k)\, p + \frac{1}{2} p^T B_k p \\ \text{s.d.} \quad & A_k p + c_k = 0 \end{aligned}$$

Es ist

$$\nabla_x L(x_k,\lambda_k)\, p = \nabla f(x_k)^T p + \left(A_k^T \lambda_k\right)^T p = \nabla f(x_k)^T p - \underbrace{\lambda_k^T c_k}_{\text{unabhängig von } x_k}.$$

3.1.1 Ungleichungsrestriktionen

$$\begin{aligned} \min \quad & f(x) \\ \text{s.d.} \quad & c_i(x) = 0 \quad i \in \mathcal{G} \\ & c_i(x) \geq 0 \quad i \in \mathcal{U} \end{aligned}$$

1. Möglichkeit: Approximation der Arbeitsmenge

   $$W_k \approx A(x_*) := \{ i \in \mathcal{U} \,|\, c_i(x) = 0 \} \cup \mathcal{G}.$$

   Also Zurückführung auf (3.1).
2. Möglichkeit:

   $$\begin{aligned} \min_p \quad & \nabla f(x_k)\, p + \frac{1}{2} p^T B_k p \\ & \nabla c_i(x_k)^T p + c_i(x_k) = 0 \qquad i \in \mathcal{G} \\ & \nabla c_i(x_k)^T p + c_i(x_k) \geq 0 \qquad i \in \mathcal{U} \end{aligned}$$

3.2 Die Wahl von B_k

3.2.1 Quasi-Newton

$$B_k \approx \nabla^2_{xx} L\left(x_k, \lambda_k\right)$$

mittels z.B. BFGS-Formel

$$\begin{aligned} B_{k+1} &= B_k - \frac{B_k \Delta x_k \Delta x_k^T B_k}{\Delta x_k B_k \Delta x_k} + \frac{y_k y_k^T}{\Delta x_k^T y_k} \\ \text{mit } y_k &= \nabla_x L\left(x_{k+1}, \lambda_{k+1}\right) - \nabla_x L\left(x_k, \lambda_k\right). \end{aligned}$$

Damit $B_{k+1} \succ 0$, braucht man $\Delta x_k^T y_k > 0$. Dies ist aber nicht immer erfüllt!

Gedämpfte BFGS Aufdatierung: Man definiert

$$y_k = \nabla_x L\left(x_{k+1}, \lambda_{k+1}\right) - \nabla_x L\left(x_k, \lambda_k\right)$$

wie zuvor und

$$\begin{aligned} r_k &= \theta_k y_k + \left(1 - \theta_k\right) B_k \Delta x_k \\ \theta_k &= \begin{cases} 1 & \Delta x_k^T y_k \geq \frac{1}{5} \nabla x_k^T B_k \nabla x_k, \\ \frac{\frac{4}{5} \Delta x_k^T B_k \Delta x_k}{\Delta x_k^T B_k \Delta x_k - \Delta x_k^T y_k} & \text{sonst} \end{cases} \in (0,1], \end{aligned}$$

also die größte Zahl ≤ 1, so daß $r_k^T \Delta x_k \geq \frac{1}{5} \Delta x_k^T B_k \Delta x_k$.

$$B_{k+1} = B_k - \frac{B_k \Delta x_k \Delta x_k^T B_k}{\Delta x_k^T B_k \Delta x_k} + \frac{r_k r_k^T}{\Delta x_k^T r_k}.$$

Dann ist $\Delta x_k^T r_k \geq \frac{1}{5} \Delta x_k^T B_k \Delta x_k > 0$, d.h. $B_{k+1} \succ 0$.

Ist $\theta_k = 1$, so ist $r_k = y_k$ (gewöhnliche BFGS). Wäre $\theta_k = 0$, so wäre $B_{k+1} = B_k$, d.h. keine Aufdatierung.

Dieses Verfahren ist nicht effektiv, wenn $\nabla^2_{xx} L\left(x_k, \lambda_k\right)$ nicht positiv definit ist!

3.2.2 Symmetrische Rank1 (SR1)-Aufdatierung

$$B_{k+1} = B_k + \frac{\left(y_k - B_k \Delta x_k\right)\left(y_k - B_k \Delta x_k\right)^T}{\left(y_k - B_k \Delta x_k\right)^T \Delta x_k^T}.$$

B_{k+1} ist nicht unbedingt positiv definit.

Gute Wahl für Trust-Region SQP-Verfahren.

3.2.3 Erweiterte (Augmented) Lagrange Funktion

Für Line-Search SQP-Verfahren benötigt man die positive Definitheit von B_k.

Ersetze die Lagrange-Funktion durch

$$L_A\left(x, \lambda; \mu\right) = f\left(x\right) + \lambda^T c\left(x\right) + \frac{1}{\mu} \left\|c\left(x\right)\right\|^2$$

Dann existiert ein μ^*, so daß $\nabla^2_{xx} L_A\left(x_k, \lambda_k; \mu\right) \succ 0$ für alle $\mu < \mu^*$.

Problem: Die Wahl von μ^*.

3.3 Schrittweiten-Bestimmung

Dafür benötigen wir eine Hilfsfunktion, z.B.

$$\varphi_1(x, \mu, \eta) = f(x) + \frac{1}{\mu} \sum_{i \in \mathcal{G}} |c_i(x)| + \frac{1}{\eta} \sum_{i \in \mathcal{U}} c_i^+(x)$$

für $\mu, \eta \in \mathbb{R}_+$, $c_i^+ := \max(0, c_i)$.

Bemerkung 3.1. φ_1 ist nicht überall differenzierbar. Aber sie ist richtungsdifferenzierbar!

Lemma 3.2. *Wenn (p_k, λ_{k+1}) eine Lösung von (3.2) ist, so gilt für die Richtungsableitung von $\varphi_1(x, \mu, \eta)$ in der Richtung p_k:*

$$\nabla(\varphi_1(x_k, \mu, \eta); p_k) \leq -p_k^T B_k p_k - \left(\mu^{-1} - \|\lambda_{k+1}\|_\infty\right) \|c(x_k)\|_1$$

Beweis. [NoWr] □

Folgerung 3.3. p_k ist eine Abstiegsrichtung für $\varphi_1(x_k, \mu)$, wenn $B_k \succ 0$ und μ hinreichend klein ist.

Die Lösung des Problemes

$$\min_{\alpha_k \in \mathbb{R}_+} \varphi_1(x_k + \alpha_k p_k, \mu, \eta)$$

liefert dann eine Schrittweite α_k.

Eine andere Hilfsfunktion ist Fletcher's erweiterte Lagrange-Funktion

$$\varphi_F = L_A(x, \lambda) = L(x, \lambda) + \frac{1}{2\mu} \|c_{\mathcal{G}}(x)\|^2 + \frac{1}{2\eta} \left\|c_{\mathcal{U}}^+(x)\right\|^2.$$

Die Hilfsfunktion ist eine weitere Variationsmöglichkeit, um verschiedene SQP-Verfahren zu erzeugen.

3.4 Line-Search SQP-Algorithmus

Algorithmus 3.4 (Line-Search SQP). Man wähle $\eta \in \left(0, \frac{1}{2}\right)$ und $\tau \in (0, 1)$.
Input: Startpunkt (x_0, λ_0) und $B_0 \approx \nabla_{xx}^2 L(x_0, \lambda_0)$.
Compute:
  Man berechne $f(x_0)$, $\nabla f(x_0)$, $c(x_0)$ und $A_0 = \nabla c(x_0)$
  for $k = 0, 1, 2, \ldots$ **do**
    if Test **then**
      STOP
    end if
    Löse (3.2) mit Voraussetzung $B_k \succ 0$ und erhalte (p_k, λ_{k+1}).
    Finde μ_k, so daß p_k eine Abstiegsrichtung für $\varphi(x, \mu)$ an der Stelle x_k ist.
    Setze $\alpha_k = 1$
    while $\varphi(x_k + \alpha_k p_k, \mu_k, \eta) > \varphi(x_k, \mu_k) + \eta \alpha_k D(\varphi(x_k, \mu_k, \eta); p_k)$ **do**
      $\alpha_k := \tau_\alpha \alpha_k$ für $\tau_\alpha \in (0, \tau)$
    end while
    $x_{k+1} := x_k + \alpha_k p_k$
    Berechne $f(x_{k+1})$, $\nabla f(x_{k+1})$, $c(x_{k+1})$, $A_{k+1} = \nabla c(x_{k+1})$
    if λ_{k+1} nicht bekannt **then**

Berechne λ_{k+1} aus $\min_\lambda \left\|\nabla f(x_{k+1}) + A_{k+1}^T \lambda\right\|_2^2$, d.h. $\lambda_{k+1} = -\left[A_{k+1}A_{k+1}^T\right]^{-1} A_{k+1}\nabla f(x_{k+1})$ ("kleinste Quadrate Methode")

end if

$\Delta x_k := x_{k+1} - x_k = \alpha_k p_k$

$y_k := \nabla_x L(x_{k+1}, \lambda_{k+1}) - \nabla_x L(x_k, \lambda_k)$

Berechne B_{k+1} aus B_k, z.B. mittels BFGS-Aufdatierung

end for

3.5 Trust-Region SQP-Verfahren

Betrachte das Problem

$$\begin{array}{ll} \min & \nabla f(x_k)\, p + \frac{1}{2} p^T B_k p \\ \text{s.d.} & A_k p + c(x_k) = 0 \\ & \|p\| \leq \Delta_k \end{array}$$

3.6 Alternativer Zugang zur Beschreibung von SQP-Verfahren

$$\begin{array}{ll} \min & f(x) \\ \text{s.d.} & c(x) \leq 0 \\ & \text{MFCQ erfüllt} \end{array}$$

Das KKT-System lautet:

$$\begin{array}{rcl} \nabla f(x) + y^T A(x) & = & 0 \\ y & \geq & 0 \\ c(x) & \leq & 0 \\ \langle y, c(x) \rangle & = & 0. \end{array}$$

Das (gestörte) Kojima-System ist

$$\begin{array}{rcl} \nabla f(x) + \left(y^+\right)^T A(x) & = & 0 \\ c(x) & = & y^- + \varepsilon y^+ \end{array}$$

mit $y_i^+ = \max\{0, y_i\}$, $y_i^- = \min\{0, y_i\}$, $\varepsilon y^+ = \left(\varepsilon_1 y_1^+, \ldots, \varepsilon_m y_m^+\right)^T$ für ein kleines $\varepsilon > 0$.

Man wende ein "verallgemeinertes" (d.h. mit verallgemeinerter Ableitung) Newton-Verfahren auf das gestörte Kojima-System an.

Durch die Wahl von ε bekommt man eine Menge von Verfahren, die "äquivalent" zur Menge der SQP-Verfahren ist. ([KuKl])

Zu einem KKT-Punkt (x, y) gehört der Punkt $(x, y + c(x))$ im Kojima-System, zu einem Kojima-Punkt (x, y) der KKT-Punkt (x, y^+).

Kapitel 4

Erweiterte Lagrange Funktionen und Methoden (Augmented Lagrangian)

Literatur: [Be]

4.1 Quadratische Straffunktion für Gleichheitsrestriktionen

$$\begin{array}{ll} \min & f(x) \\ \text{s.d.} & h(x) = 0 \in \mathbb{R}^m \end{array}$$

Für beliebigen Strafparameter $\mu > 0$ betrachte das Problem

$$\min_{x \in \mathbb{R}^n} Q(x;\mu) \equiv f(x) + \frac{1}{2}\mu \, \|h(x)\|_2^2$$

Standardvoraussetzung:

$$f, h \in C^{1,1}(\mathbb{R}^n) \qquad \lim_{\|x\| \to \infty} f(x) = \infty.$$

Satz 4.1. *Die Standardvoraussetzung impliziert für jede Folge* $\mu_k \to \infty$ *die Existenz einer Folge von globalen Minima*

$$x_k \equiv \operatorname{argmin}\{Q(x, \mu_k)\}$$

Diese hat mindestens einen Häufungspunkt $x_* \in \mathbb{R}^n$ *und es gilt*

$$x_* \in \operatorname{argmin}\{f(x) \,|\, x \in \mathbb{R}^n, h(x) = 0\}.$$

Ist an x_* *die LICQ Bedingung* $\operatorname{rank}(\nabla h(x_*)) = m$ *erfüllt und sind* $\lambda_* \in \mathbb{R}^m$ *die entsprechenden Lagrange-Multiplikatoren, so gilt*

$$\lim_{k \to \infty} \mu_k h(x_*) = \lambda_* \in \mathbb{R}^m.$$

Wobei o.B.d.A. angenommen wurde, daß nur ein Häufungspunkt existiert.

Beweis. Wegen $Q(x,\mu) \geq f(x)$ gilt $\lim_{\|x\|\to\infty} Q(x,\mu) = \infty$, es existieren also immer globale Minima $x_k \in \operatorname{argmin}\{Q(x,\mu_k)\}$. Die Folge ist beschränkt, hat also einen Häufungspunkt x_* (o.B.d.A. eindeutig).

$$f_* = \inf_{x\in\mathbb{R}^n} f(x) \leq f(x_k) \leq Q(x_k,\mu_k) \leq Q(\tilde{x}_*,\mu_k) = f(\tilde{x}_*) < \infty,$$

wobei $\tilde{x}_*$ ein globaler Minimalpunkt von f ist. Daraus folgt im Grenzwert

$$\lim_{k\to\infty} \frac{1}{\mu_k}\left[Q(x_k,\mu_k) - f(x_k)\right] \leq \lim_{k\to\infty} \frac{1}{\mu_k}\left(f(\tilde{x}_*) - f_*\right) = 0$$

Wegen

$$\lim_{k\to\infty} \frac{1}{\mu_k}\left[Q(x_k,\mu_k) - f(x_k)\right] = \lim_{k\to\infty} \frac{1}{2}\|h(x_k)\|^2 = \frac{1}{2}\|h(x_*)\|^2$$

folgt

$$h(x_*) = 0.$$

Also ist der Häufungspunkt zulässig und wegen

$$\lim_{k\to\infty} f(x_k) = f(x_*) \leq f(\tilde{x}_*)$$

auch global optimal.

Zum Beweis der letzten Aussage nutze die Optimalitätsbedingung

$$\nabla_x Q(x_k,\mu_k) = \nabla f(x_k) + \mu_k \nabla h(x_k)^T h(x_k) = 0.$$

Wegen LICQ hat $A_k := \nabla h(x_k) \to \nabla h(x_*)$ für alle großen k vollen Rank hat, so daß $A_k A_k^T \in \mathbb{R}^{n\times n}$ regulär ist.

$$\begin{aligned} A_k \nabla f(x_k) + A_k A_k^T \mu_k h(x_k) &= 0 \\ \Rightarrow \mu_k h(x_k) &= -\left(A_k A_k^T\right)^{-1} A_k \nabla f(x_k) \\ \xrightarrow{k\to\infty} \quad \lambda_* &= \left(A_* A_*^T\right)^{-1} A_* \nabla f(x_*) \end{aligned}$$

ist eindeutige Lsg. von KKT-Bedingung $\nabla f(x_*) + A(x_*)^T \lambda_* = 0$ □

Vorteil: Einfache Anwendung eines unbeschränkten Optimierungsverfahrens auf ein beschränktes Problem. Die $Q(x,\mu)$ sind im Gegensatz zu L_1-Straffunktionen so oft diff.bar wie f und h selbst.

Nachteil: Mit wachsenden μ werden die unbeschränkten Probleme $Q(x,\mu)$ immer schwerer lösbar, insbesondere geht die Konditionszahl der Hessematrix

$$\nabla_x^2 Q(x_*;\mu) = \nabla^2 f(x_*) + \mu \nabla h(x_*)^T \nabla h(x_*) + \sum_{i=1}^{n} \lambda_*^{(i)} \nabla^2 h(x_*)$$

mit μ gegen unendlich.

Es sei denn, es gilt $m = n$, was im allgemeinen nicht von Interesse ist.

Bemerkung 4.2. Der Satz gilt im "wesentlichen" schon dann, wenn die Teilprobleme so genau gelöst werden, daß

$$\lim_{k\to\infty} \|\nabla_x Q(x_k,\mu_k)\| \equiv \lim_{k\to\infty} \left\|\nabla f(x_k) + \mu \nabla h(x_k)^T h(x_k)\right\| = 0.$$

4.2 Lagrange Multiplikatoren Methode (1. Ordnung)

Idee: Vermeide $\mu_k \to \infty$ durch Addition des Lagrange-Termes $\lambda^T h(x)$. Dies ergibt die erweiterte Lagrange Funktion

$$E(x,\lambda,\mu) \equiv f(x) + \lambda^T h(x) + \frac{1}{2}\mu \|h(x)\|^2 .$$

Satz 4.3. *Voraussetzung: $f, h \in C^{2,1}(\mathbb{R}^n)$ und $x_* \in h^{-1}(0)$ ist lokales Minimum an dem die LICQ und die hinreichenden lokalen Optimalitätsbedingungen erster und zweiter Ordnung erfüllt sind.*

Behauptung: Für alle hinreichend großen μ ist x_* dann unbeschränktes lokales Minimum von $E(x,\lambda_*,\mu)$, wobei $\lambda_* \in \mathbb{R}^m$ die exakten Lagrangemultiplikatoren an x_* sind.

Beweis. Optimalitätsbedingung 1. Ordnung:

$$\nabla_x E(x_*,\lambda_*,\mu) = \nabla f(x_*) + \nabla h(x_*)^T \lambda_* + \mu \underbrace{\nabla h(x_*)^T h(x_*)}_{=0} = \nabla_x L(x_*,\lambda_*) = 0$$

Optimalitätsbedingung 2. Ordnung: Es ist zu zeigen, daß für hinreichend große μ

$$\nabla_x^2 E(x_*,\lambda_*,\mu) = \nabla_x^2 L(x_*,\lambda_*) + \mu \nabla h(x_*)^T \nabla h(x_*) \succ 0$$

gilt. Mit $A = \nabla h(x_*)$ und $Z \in \mathbb{R}^{(n-m)\times n}$, so daß $AZ^T = 0$ und $ZZ^T = I \in \mathbb{R}^{(n-m)\times(n-m)}$ läßt sich jeder Vektor u in $\mathbb{R}^n$ darstellen als

$$u = Z^T v + A^T w.$$

Multiplikation von $\nabla_x^2 E(x_*,\lambda_*,\mu)$ von rechts und links mit u ergibt

$$\begin{aligned}
u^T \nabla_x^2 E(x_*,\lambda_*,\mu)\, u &= \left(v^T Z + w^T A\right)\left(\nabla_x^2 L(x_*,\lambda_*) + \mu A^T A\right)\left(Z^T v + A^T w\right) \\
&= v^T Z \nabla_x^2 L(x_*,\lambda_*) Z^T v + \mu w^T \left(AA^T\right)^2 w + 2 w^T A \nabla_x^2 L(x_*,\lambda_*) Z^T v \\
&\quad + w^T A \nabla_x^2 L(x_*,\lambda_*) A^T w \\
&\geq \underbrace{\alpha \|v\|^2}_{\substack{\text{mit } \alpha>0 \text{ wegen SSC=} \\ \text{2. Order Sufficiency Condition}}} + \underbrace{\beta\mu \|w\|^2}_{\substack{\text{mit } \beta>0 \\ \text{wegen LICQ}}} - \underbrace{2\gamma \|w\|\,\|v\|}_{\text{mit } \gamma<\infty} - \underbrace{\delta \|w\|^2}_{\text{mit } \delta<\infty} \\
&= \alpha \|v\|^2 - 2\gamma \|w\|\,\|v\| + (\beta\mu - \delta)\|w\|^2 \\
&= (\|v\|, \|w\|)^T \begin{pmatrix} \alpha & -\gamma \\ -\gamma & \beta\mu - \delta \end{pmatrix} \begin{pmatrix} \|v\| \\ \|w\| \end{pmatrix}.
\end{aligned}$$

Die Matrix hat α als positives diagonales Element, ist also positiv definit, gdw. ihre Determinante $\alpha(\beta\mu - \delta) - \gamma^2$ positiv ist. Dies ist sicherlich für hinreichend große μ der Fall. □

Bemerkung 4.4. Entsprechend hat $E(x,\lambda,\mu)$ auch für alle $\lambda \approx \lambda_*$ und μ hinreichend groß ein lokales Minimum nahe x_* (Im allgemeinen nicht in $h^{-1}(0)$, d.h. nicht zulässig bzgl. $h(x) = 0$).

Frage: Wie kann man die λ_k justieren, so daß im Grenzwert $\lambda_k \to \lambda_*$ gilt?

Antwort: Multiplikatorenregel 1. Ordnung: Wegen

$$x_{k+1} = \text{argmin}\left\{E(x,\lambda_k,\mu_k)\right\}$$

ist

$$\begin{aligned}
0 &= \nabla f(x_{k+1}) + \nabla h(x_{k+1})^T \lambda_k + \nabla h(x_{k+1})^T h(x_{k+1})\, \mu_k \\
&= \nabla f(x_{k+1}) + \nabla h(x_{k+1})^T \underbrace{\left(\lambda_k + h(x_{k+1})\, \mu_k\right)}_{\equiv \lambda_{k+1}}.
\end{aligned}$$

Dies ergibt die (natürliche) Aufdatierung

$$\lambda_{k+1} = \lambda_k + \mu_k h(x_{k+1}).$$

Außerdem sollte $\mu_{k+1} \geq \mu_k$ möglichst so gewählt werden, daß $\nabla_x^2 E(x_{k+1}, \lambda_{k+1}, \mu_{k+1})$ positiv definit ist.

Satz 4.5. *Unter den Voraussetzungen von Satz 4.3 folgt aus μ hinreichend groß und $\|x_0 - x_*\|$, $\|\lambda_0 - \lambda_*\|$ hinreichend klein die lineare Konvergenz der Iteration*

$$\begin{aligned} x_{k+1} &= \operatorname*{argmin}_x \{E(x, \lambda_k, \mu_k)\}, \\ \lambda_{k+1} &= \lambda_k + h(x_k)\mu \end{aligned}$$

zum beschränkten Minimum x_* mit Multiplikatoren λ_*.

Beweis. Benutze Standardergebnis für Iteration $z = G(y)$ mit $y_* = G(y_*)$ und $G \in C^{1,1}(\mathbb{R}^n)$: Falls alle Eigenwerte von $G_y|_{y_*} \equiv \frac{dz}{dy}|_{y_*} \in \mathbb{R}^{(n+m)\times(n+m)}$ betragsmäßig kleiner eins sind, folgt aus $\|y_0 - y_*\| \approx 0$, daß $y_{k+1} = G(y_k) \xrightarrow{k\to\infty} y_*$.

Hier haben wir eine implizite Beziehnung zwischen $y = (x_k, \lambda_k)$ und $z = (x_{k+1}, \lambda_{k+1})$, nämlich

$$0 \equiv F(y,z) = F((x_k, \lambda_k), (x_{k+1}, \lambda_{k+1})) := \begin{bmatrix} \nabla_x L(x_{k+1}, \lambda_k) + \mu \nabla h(x_{k+1})^T h(x_{k+1}) \\ \lambda_{k+1} - \lambda_k - \mu h(x_{k+1}) \end{bmatrix}.$$

Daraus folgt

$$\begin{aligned} F_y = \frac{\partial F}{\partial(x_k, \lambda_k)} &= \begin{bmatrix} 0 & \nabla h(x_{k+1})^T \\ 0 & -I \end{bmatrix} \\ F_z = \frac{\partial F}{\partial(x_{k+1}, \lambda_{k+1})} &= \begin{bmatrix} \nabla_x^2 E(x_{k+1}, \lambda_{k+1}) & 0 \\ -\mu \nabla h(x_{k+1}) & I \end{bmatrix}. \end{aligned}$$

ρ ist Eigenwert von $\frac{dz}{dy} = F_z^{-1} F_y$, gdw.

$$\begin{aligned} & 0 = \det\left[\rho I + F_z^{-1} F_y\right] \\ \Leftrightarrow\quad & 0 = \det(\rho F_z + F_y) \qquad \text{da } \det F_z \neq 0 \\ \Leftrightarrow\quad & \rho = 0 \text{ oder } P(\rho) := \det\left(\begin{bmatrix} \nabla_x^2 E(x_{k+1}, \lambda_{k+1}) & \frac{1}{\rho}\nabla h^T(x_{k+1}) \\ -\mu \nabla h(x_{k+1}) & \left(1 - \frac{1}{\rho}\right) I \end{bmatrix}\right) = 0. \end{aligned}$$

Zu zeigen: Alle solche ρ sind reell und liegen für hinreichend großes $\mu > 0$ im Inneren des Intervalles $(-1, 1)$.

Ab jetzt werden alle Ableitungen direkt an der Lösung (x_*, λ_*) ausgewertet. Die hergeleiteten Ungleichheitsaussagen gelten dann aufgrund der Stetigkeit in einer Umgebung.

Fall $\rho = 1$: Dann ist

$$P(1) = \det\left(\begin{bmatrix} \nabla_x^2 L + \mu \nabla h^T \nabla h & \nabla h^T \\ \nabla h & 0 \end{bmatrix}\right)(-\mu)^m.$$

Wegen der Optimalitätsbedingung 2. Ordnung gilt für alle Z mit $\nabla h\, Z = 0$

$$Z \nabla_x^2 L Z^T = Z\left(\nabla_x^2 L + \mu \nabla h^T \nabla h\right) Z^T \succ 0.$$

Daraus folgt wiederum $P(1) \neq 0$.

Fall $\rho \neq 1$: Dann ist

$$\begin{aligned}
P(\rho)\frac{\rho^m}{(\rho-1)^m} &= \det\left(\begin{bmatrix} \nabla_x^2 E & \frac{1}{\rho-1}\nabla h^T \\ -\mu\nabla h & I \end{bmatrix}\right) \\
&= \det\left(\begin{bmatrix} \nabla_x^2 E + \frac{\mu}{\rho-1}\nabla h^T\nabla h & 0 \\ 0 & I \end{bmatrix}\right) \\
&= \det\left(\nabla_x^2 L + \left(\mu + \frac{\mu}{\rho-1}\right)\nabla h^T\nabla h\right) \\
&= \det\left(\nabla_x^2 L + \frac{\mu\rho}{\rho-1}\nabla h^T\nabla h\right).
\end{aligned}$$

Da die Matrix symmetrisch ist, sind alle Wurzeln ρ reell. Außerdem wurde im Beweis von Satz 4.3 gezeigt, daß die Matrix positiv definit ist, falls $\frac{\mu\rho}{\rho-1} \geq \bar{\mu}$ für eine untere Schranke $\bar{\mu}$. Für $\mu \geq \bar{\mu}$ und $\rho > 1$ oder $\mu \geq 2\bar{\mu}$ und $\rho \leq -1$ ist dies erfüllt. □

4.3 Verallgemeinerung auf Ungleichungen

$g_j(x) \leq 0$ gdw. $g_j(x) + s_j^2 = 0$ für eine Schlupfvariable s_j^2.

Dies ergibt in der erweiterten Lagrange-Funktion den Term

$$E_j = \lambda_j\left(g_j(x) + s_j^2\right) + \frac{\mu}{2}\left(g_j(x) + s_j^2\right)^2,$$

wobei dafür Sorge getragen wird, daß immer $\lambda_j \geq 0$ gilt.

E_j hat bzgl. $g_j(x) + s_j^2$ das unbeschränkte Minimum

$$g_j(x) + s_j^2 = -\frac{\lambda_j}{\mu}.$$

Dies kann nur durch die Wahl von s_j erreicht werden, falls $g_j(x) \leq -\frac{\lambda_j}{\mu}$. Dann ist $E_j \equiv -\frac{\lambda_j^2}{2\mu}$. Ist $g_j(x) > -\frac{\lambda_j}{\mu}$, so ist $E_j = -\lambda_j g_j(x) + \frac{\mu}{2}g_j^2(x)$.

Für $\mu \to \infty$ erhält man $E_j \to \begin{cases} 0 & \text{für } g_j \leq 0 \\ \infty & \text{für } g_j > 0 \end{cases}$.

E_j wird auch "*shifted quadratic penalty function*" genannt.

E_j ist bzgl. $g_j$und damit x zwar stetig differenzierbar, hat aber am "Schaltpunkt" $g_j = -\frac{\lambda_j}{\mu}$ einen Sprung in der zweiten Ableitung.

Bemerkung 4.6. Während die Slackvariable s_j^2 explizit eliminiert wurde, bleibt der Lagrange Multiplikator λ_j erhalten. Die Aufdatierungsformel 1. Ordnung ist einfach

$$\lambda_j^{(n+1)} = \max\left\{0, \lambda_j^{(n)} + \mu^{(n)} g_j\left(x^{(n)}\right)\right\}.$$

Anhang A

Interior Point Methods

Vorlesung von Prof. Florian A. Potra, University of Maryland Baltimore Country.

Dantzig, 1947: Simplex Algorithm

$$(P) \qquad \begin{array}{rl} \min & c^T x \\ \text{s.t.} & Ax = b \\ & x \geq 0 \end{array} \qquad\qquad (D) \quad \begin{array}{rl} \max & b^T y \\ \text{s.t.} & A^T y + s = c \\ & s \geq 0 \end{array}$$

$$\begin{aligned} \mathcal{P} &:= \{x \in \mathbb{R}^n \,|Ax = b, x \geq 0\} \\ \mathcal{D} &:= \left\{(s,y) \in \mathbb{R}^m \,\middle|A^T y + s = c, s \geq 0\right\} \end{aligned}$$

x primal variables, y dual variables, s slack variables

We observe, that

$$c^T x - b^T y = x^T c - y^T b = x^T \left(A^T y + s\right) - y^T A x = x^T A^T y + x^T s - y^T A x = x^T s \geq 0$$

So, for all $x \in \mathcal{P}$ and all $(s,y) \in \mathcal{D}$

$$0 \leq c^T x + b^T y = x^T s =: \text{duality gap}$$

Satz A.1. *$\exists x^* \in \mathcal{P}$, $(s^*, y^*) \in \mathcal{D}$, s.t. ${x^*}^T s^* = 0 \Rightarrow c^T x^* = \min_{x \in \mathcal{P}} c^T x$, $b^T y^* = \max_{(y,s) \in \mathcal{D}} b^T y$*

Beweis.

$$c^T x \quad \geq \quad b^T y \qquad \forall x \in \mathcal{P}, (s,y) \in \mathcal{D}$$

So,

$$c^T x \geq b^T y^* = c^T x^* \qquad \forall x \in \mathcal{P}$$

on the one hand, and

$$b^T y \leq c^T x^* = b^T y^* \qquad \forall\, (s,y) \in \mathcal{D}. \qquad \square$$

And we see,

$$\begin{aligned} 0 &\leq c^T x - c^T x^* \leq x^T s \qquad x \geq 0 \\ 0 &\leq b^T y^* - b^T y \leq x^T s \qquad s \geq 0 \end{aligned}$$

$x^T s = 0 \Leftrightarrow x_i s_i = 0 \;\forall i \in \{1, \ldots, n\}$. (complementarity)

$$(PD)\qquad \begin{array}{rcl} xs &=& 0\\ Ax-b &=& 0\\ A^Ty+s-c &=& 0\\ x &\geq& 0\\ s &\geq& 0 \end{array}$$

Denote by

$$\begin{array}{rcl} z &=& \begin{pmatrix} x\\ s\\ y \end{pmatrix}\\ F(z) &=& \begin{pmatrix} xs\\ Ax-b\\ A^Ty+s-c \end{pmatrix} \qquad F:\mathbb{R}^{2n+m}\to\mathbb{R}^{2n+m}. \end{array}$$

We need to find a point z with $F(z)=0$. Using Newton's Method, we obtain

$$\begin{array}{rcl} z^+ &=& z-F'(z)^{-1}F(z)\\ F'(z) &=& \begin{pmatrix} S & X & 0\\ A & 0 & 0\\ 0 & I & A^T \end{pmatrix} \qquad S=\begin{pmatrix} s_1 & & 0\\ & \ddots & \\ 0 & & s_n \end{pmatrix} \end{array}$$

Let A have full rank.

Annahme: $\exists x^0\in\mathcal{P},\ (s^0,y^0)\in\mathcal{D},\ x^0>0,\ s^0>0$ (*interior point condition*)

Now, denote by

$$F_\tau(z) := \begin{pmatrix} xs-\tau e\\ Ax-b\\ A^Ty+s-c \end{pmatrix}, \qquad e=\begin{pmatrix} 1\\ \vdots\\ 1 \end{pmatrix}\in\mathbb{R}^n,$$

so $F_0(z)=F(z)$.

The idea of path following interior point method is: Given a point

$$z\in\mathcal{F}^0=\left\{ z=\begin{pmatrix} x\\ s\\ y \end{pmatrix} \middle| Ax=b, A^Ty+s=c, x>0, y>0 \right\}$$

compute $\mu=\frac{x^Ts}{n}$ and apply a newton step to $F_{\bar{\mu}}(t)=0$ where $\bar{\mu}=(1-\gamma)\mu$ for some $\gamma\in(0,1)$.

$$\begin{array}{rl} & \Delta z=\begin{pmatrix} u\\ v\\ w \end{pmatrix}=-F_\tau'(z)^{-1}F_\tau(z)\\ \Leftrightarrow & F_\tau'(z)\Delta z=-F_\tau(z)\\ \Leftrightarrow & \begin{array}{rcl} Su+Xv &=& \tau e-Xs\\ Au &=& b-Ax=0\\ v+A^Tw &=& c-s-A^Ty=0 \end{array} \end{array} \tag{A.1}$$

Lemma A.2. *If $z\in\mathcal{F}^0$ then the solution of the system (A.1) satisfies*

1. $u^Tv=0$

2. $\|Du\|^2 + \|D^{-1}v\|^2 = \left\|(XS)^{-\frac{1}{2}}(\tau e - xs)\right\|^2 =: \|a\|^2$, $\sqrt{x} = \left(\sqrt{x_1}, \ldots, \sqrt{x_n}\right)^T$

3. $\|uv\| \leq \frac{1}{\sqrt{8}}\|a\|^2$

4. $\|uv\|_\infty \leq \frac{1}{4}\|a\|^2$

5. $\|uv\|_1 \leq \frac{1}{2}\|a\|^2$

where $D = X^{-\frac{1}{2}}S^{\frac{1}{2}}$.

Beweis.

1. $Au = 0 \Leftrightarrow u \in \ker(A)$ and $v + A^T w = 0 \Leftrightarrow v \in \mathrm{Range}\left(A^T\right)$
   So, $u^T v = 0$, because $u^T v = \langle u, v\rangle = \left\langle u, -A^T w\right\rangle = \langle Au, -w\rangle = 0$.
2. $Su + Xv = \tau e - xs$
   $\bar{u} = Du$, $\bar{v} = D^{-1}v$, $a = \bar{u} + \bar{v}$
   $Du + D^{-1}v = a$,

$$\|a\|^2 = \left\|Du + D^{-1}v\right\|^2 = \|Du\|^2 + \left\|D^{-1}v\right\|^2 + 2\,(Du)^T\,(Dv)^{-1} = 2u^T v = 0 \qquad \square$$

Satz A.3. *Assume* $z \in \mathcal{F}^0$, $\|xs - \mu e\| \leq \alpha\mu$, $\alpha \in (0,1)$, $\mu = \frac{x^T s}{n}$. *Then by taking* $\gamma = \ldots$ *and*

$$\begin{aligned} \bar{\mu} &= \tau = (1-\gamma)\,\mu, \\ \bar{z} &= z - \left(F_\tau'(z)\right)^{-1} F(z), \end{aligned}$$

we have $\bar{z} \in \mathcal{F}^0$ *and*

$$\left\|\bar{x}^T\bar{s} - \bar{\mu}e\right\| \leq \alpha\bar{\mu}.$$

Beweis. $z(\gamma) = \bar{z}$,

$$\begin{aligned} \bar{x}\bar{s} &= (x+u)(s+v) = xs + su + xv + uv \\ &= xs + \tau e - xs + uv \qquad \text{by (A.1)} \\ &= \tau e + uv \\ &= (1-\gamma)\,\mu e + uv \\ \Rightarrow \bar{\mu} &= \frac{\bar{x}^T\bar{s}}{n} = (1-\gamma)\,\mu, \qquad \text{because } \sum \bar{x}_i\bar{s}_i = n\,(1-\gamma)\,\mu + \underbrace{\sum u_i v_i}_{=0} \\ \Rightarrow \bar{x}\bar{s} - \bar{\mu}e &= uv \end{aligned}$$

so $\|\bar{x}\bar{s} - \bar{\mu}e\| \leq \alpha\bar{\mu} \Leftrightarrow \|uv\| \leq \alpha\bar{\mu}$.

From the Lemma: $\|uv\| \leq \frac{1}{\sqrt{8}}\left\|(XS)^{-\frac{1}{2}}\left((1-\gamma)\,\mu e - xs\right)\right\|^2$, so we have to determine γ such that $\left\|(XS)^{-\frac{1}{2}}\left((1-\gamma)\,\mu e - xs\right)\right\| \leq \alpha\bar{\mu} = \alpha\,(1-\gamma)\,\mu$.

We have

$$\begin{aligned}
\|xs-\mu e\|^2 &= \sum (x_i s_i - \mu)^2 \le \alpha^2\mu^2 \\
\Rightarrow |x_i s_i - \mu| &\le \alpha\mu \\
\Rightarrow x_i s_i &\ge (1-\alpha)\mu \\
\Rightarrow \left\|(XS)^{-\frac{1}{2}}\left((1-\gamma)\mu e - xs\right)\right\|^2 &\le \frac{1}{\min x_i s_i}\|(1-\gamma)\mu e - xs\|^2 \\
&\le \frac{\|\mu e - xs\|^2 + \|(1-\gamma)\mu e - \mu e\|^2}{(1-\alpha)\mu} \\
&\le \frac{1}{(1-\alpha)\mu}\left(\alpha^2\mu^2 + \mu^2\gamma^2 n\right) \\
&= \frac{\alpha^2+\gamma^2 n}{(1-\alpha)}\mu \\
\Rightarrow \|\bar{x}\bar{s} - \bar{\mu}e\| &\le \frac{1}{\sqrt{8}}\cdot\frac{\alpha^2+\gamma^2 n}{(1-\alpha)}\mu \\
&\overset{?}{\le} \alpha(1-\gamma)\mu
\end{aligned}$$

Assume the stronger condition $\alpha \le \frac{\sqrt{8}}{1+\sqrt{8}}$. So $\alpha^2\left(1+\sqrt{8}\right) < \sqrt{8}\alpha$ implies $\alpha^2 < \sqrt{8}\alpha(1-\alpha)$.

Then, $\frac{\alpha^2+\gamma^2 n}{(1-\alpha)\sqrt{8}}\mu \le \alpha(1-\gamma)\mu$ is satisfied, iff

$$\begin{aligned}
\alpha^2+\gamma^2 n &\le \sqrt{8}\alpha(1-\alpha)(1-\gamma) \\
\Leftrightarrow \gamma^2 n + \underbrace{\sqrt{8}\alpha(1-\alpha)}_{2\lambda}\gamma - \underbrace{\left(\sqrt{8}\alpha(1-\alpha)-\alpha^2\right)}_{\sigma>0} &\le 0 \\
\gamma^* &= \frac{-\lambda+\sqrt{\lambda^2+\sigma n}}{n} \\
&= \frac{\sigma n}{n\left(\sqrt{\lambda^2+\sigma n}+\lambda\right)} \\
&\xrightarrow{n\to\infty} \frac{\sqrt{\sigma}}{\sqrt{n}}.
\end{aligned}$$

Want an elegant lower bound for γ^*. Set $\gamma^* = \frac{\delta}{\sqrt{n}}$,

$$\delta = \frac{\sigma}{\sqrt{\frac{\lambda^2}{n}+\sigma}+\frac{\lambda}{\sqrt{n}}} \ge \frac{\sigma}{\sqrt{\frac{\lambda^2}{2}+\sigma}+\frac{\lambda}{\sqrt{2}}} =: \hat{\delta}$$

If we take $\hat{\gamma} = \frac{\hat{\delta}}{\sqrt{n}}$ we have $\|\bar{x}\bar{s}-\bar{\mu}e\| \le \alpha\bar{\mu}$ provided $\alpha < \frac{\sqrt{8}}{1+\sqrt{8}}$. □

Algorithmus A.4.
Input: $z^0 \in \mathcal{F}^0$, $z^0 \in N_\alpha$ where $N_\alpha = \{z \,|\|xs-\mu e\| \le \alpha\}$
Compute:
  for $k = 0, 1, \ldots$ **do**
    $z \leftarrow z^k$
    determine γ such that $\bar{z} \in N_\alpha$
    $z^{k+1} \leftarrow \bar{z}$, $\gamma_k \leftarrow \gamma$
  end for

Properties: $z^k \in N_\alpha$, $\mu_{k+1} \le (1-\gamma_k)\mu_k$, $k = 01, 2, \ldots$

Korollar A.5. $\mu_k \leq \left(1 - \frac{\hat{\delta}}{\sqrt{n}}\right)^k \mu_0$

$\frac{\mu_k}{\mu_0} \leq \varepsilon$ for $\left(1 - \frac{\hat{\delta}}{\sqrt{n}}\right)^k \leq \varepsilon$, i.e. $k \log\left(1 - \frac{\hat{\delta}}{\sqrt{n}}\right) \leq \log \varepsilon$.

Bemerkung A.6. If A, b, c are integers and L ist the binary length of data and $\varepsilon < 2^{-L}$ then z can be rounded to an exact solution of LP in $O\left(n^3\right)$ operations.

An exact solution of (PD) is obtained in $O\left(\sqrt{n}L\right)$ iterations + $O\left(n^3\right)$ arithmetic operations.

Smart algorithms determine γ_k such that $\gamma_k \nearrow 1$.

From $su + xv = (1-\gamma)\mu e - xs$ obtain the

- affine scaling direction from
$$su^a + xv^a = -xs \tag{A.2}$$
Newtonstep: $-F'(z)^{-1} F_0(z)$
- centering scaling direction from
$$su^c + xv^c = \mu e - xs \tag{A.3}$$
Newtonstep: $-F'(z)^{-1} F_\mu(z)$

With $u = \gamma u^a + (1-\gamma) u^c$, $v = \gamma v^a + (1-\gamma) v^c$, it holds

$$\begin{aligned} su + xv &= \gamma\left(su^a + xv^a\right) + (1-\gamma)\left(su^c + xv^c\right) \\ &= \gamma(-xs) + (1-\gamma)(\mu e - xs) \\ &= (1-\gamma)\mu e - xs \end{aligned}$$

$\|uv\| \leq \alpha(1-\gamma)\mu$ holds, if

$$\left\|\gamma^2 u^a v^a + \gamma(1-\gamma)\left(u^a v^c + u^c v^a\right) + (1-\gamma)^2 u^c v^c\right\| \leq \alpha(1-\gamma)\mu \tag{A.4}$$

Solve (A.4) for γ.

Algorithmus A.7 (Gonzaga, Largest Step). Solve (A.2) and (A.3) and then get the longest γ, that satisfies (A.4).

Note: (A.2) and (A.3) have the same Jacobian.

$$\begin{aligned} N_\alpha &= \left\{ z \,\middle|\, \left\| \frac{xs}{\mu} - e \right\|_2 \leq \alpha \right\} \\ N_\alpha^- &= \left\{ z \,\middle|\, \left\| \left(\frac{xs}{\mu} - e \right)^- \right\|_\infty \leq \alpha \right\} \\ \mathcal{D}_\beta &= N_{(1-\beta)}^- = \{ z \,|\, xs \geq \beta\mu e \} \\ \lim_{\alpha \nearrow 1} N_\alpha^- &= \mathcal{F} \end{aligned}$$

Literaturverzeichnis

[JaSt] Jarre, Stoer, Optimierung (gründlich)

[NoWr] Nocedal, Wright, Numerical Optimization (lesbar)

[KiSc] Kielbasinski, Schwetlick, Numerische lineare Algebra

[KuKl] Kummer, Klatte, Nonsmooth Equations in Optimization

[Be] D. Bertsehas, Constraint Optimization and Lagrange Multiplier Methods, 1982

Index