

Zwischenfolgerungen (für den unrestringierten Fall)

- ▶ Gradienten kosten nur ein kleines Vielfaches der zu Grunde liegenden Funktionsauswertung vorausgesetzt diese ist durch einen Auswertungscode gegeben.
- ▶ Hesse- und Jacobimatrizen, e.g. $\nabla^2 f$ und im beschränkten Falle $\nabla h, \nabla c$, können sehr teuer zu faktorisieren sein, falls sie keine geeignete Dünnbesetztheitsstruktur besitzen.
- ▶ Gradientenbasierte quasi-Newton Methoden sind ein guter Kompromiss zwischen langsamen ableitungsfreien Verfahren und teuren Methoden zweiter Ordnung wie z.B. Newton.
- ▶ Bei unrestringierten Problemen kann durch Strahlsuche Konvergenz zu einem stationären Punkt erzwungen werden.
- ▶ Globale Optimierung ist extrem teuer und/oder sehr unzuverlässig.

KKT Optimalitätsbedingungen für restringierte Minima

An lokalen Minimalpunkten muss die Lagrangefunktion

$$L(x, \lambda, \mu) = f(x) + \sum_{j=1}^m \lambda_j h_j(x) + \sum_{i=1}^n \mu_i c_i(x) \quad \text{mit } \mu_i \geq 0$$

nach Karush (1939) und Kuhn-Tucker (1951) die folgenden Bedingungen erster Ordnung erfüllen

$$\begin{aligned} \nabla_x L(x, \lambda, \mu) &= 0, & \nabla_\lambda L(x, \lambda, \mu) &= 0 \\ \nabla_\mu L(x, \lambda, \mu) &\leq 0 \leq \mu, & \mu^T \nabla_\mu L(x, \lambda, \mu) &= 0 \end{aligned}$$

Als Bedingung zweiter Ordnung muss gelten für beliebige $v \in \mathbb{R}^n$

$$\nabla h(x)v = 0, \quad \text{diag}(\mu_i) \nabla c(x)v = 0 \implies v^T \nabla_x^2 L(x, \lambda, \mu)v \geq 0$$

Restringierte Nichtlineare Optimierung

Umformungstricks zu unrestringiertem Problem

Ungleichungsrestriktion *konvertierbar zu* Gleichheitsrestriktion
 Vorzeichenbedingung
 oder quadrierter Schlupf

$$c_j(x) \leq 0 \iff c_j(x) + z_j^e = 0 \quad \text{mit } e = 1 \text{ und } z_j \geq 0 \quad \text{oder } e = 2$$

Bewertungsfunktion:

$$f_\rho(x) \equiv f(x) + \rho p(h(x)) + \frac{1}{\rho} b(c(x))$$

wobei

$$\begin{aligned} p(z) &= \text{Strafe für } z \neq 0 \quad \text{e.g. } p(z) = z^2 \\ b(z) &= \text{Barriere für } z \rightarrow 0 \quad \text{e.g. } b(z) = -\ln(z) \end{aligned}$$

Unter ziemlich allgemeinen Voraussetzungen gilt dann

$$\lim_{\rho \rightarrow 0} x_\rho \equiv \operatorname{argmin}_{x \in \mathbb{R}^n} f_\rho(x) = x_* \in \operatorname{argmin}(f|M)$$

Erweiterte Lagrange Methoden (MINOS(1970), LANCELOT(1988))

Minimiere

$$L_\rho(x, \lambda, \mu) = L(x, \lambda, \mu) + \frac{1}{2} \rho (\|h\|^2 + \sum_{i=1}^p \max(0, c_i)^2)$$

welche identische Minima $x_\rho = x_*$ für grosses ρ und korrektes λ, μ hat.

Sequentielle Quadratische Programmierung (Wilson(1963), Powell)

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

$$\begin{aligned} \text{s.t. } h(x) + \nabla h(x)s &= 0 & \text{mit linearen Restriktionen} \\ c(x) + \nabla c(x)s &\leq 0 \end{aligned}$$

wobei

$$B \approx \nabla_x^2 L(x, \lambda, \mu) = \nabla^2 f(x) + \sum_{j=1}^m \lambda_j \nabla^2 h_j(x) + \sum_{i=1}^p \mu_i \nabla^2 g_i(x)$$

alle Krümmungsinformationen enthält.

problem	nv	nc	ipop	knitro	-d	-a	loqo	pennon	snopt
<i>bearing</i> ₂₀₀	40000	0	10	236	10	93	11	29	fail
<i>bearing</i> ₄₀₀	160000	0	106	856	67	1018	89	173	fail
<i>camshape</i> ₁₆₀₀	1600	3200	3	31	1	c	4	fail	10
<i>camshape</i> ₆₄₀₀	6400	12800	17	178	7	t	i	fail	177
<i>elec</i> ₂₀₀	600	200	84	14	22	27	67	34	46
<i>elec</i> ₄₀₀	1200	400	994	63	164	190	1296	183	78
<i>gasoil</i> ₁₆₀₀	16001	15998	8	164	120	923	15	t	225
<i>gasoil</i> ₃₂₀₀	32001	31998	26	1239	748	5163	110	t	688
<i>marine</i> ₈₀₀	19215	19192	8	11	14	3039	28	t	t
<i>marine</i> ₁₆₀₀	38415	38392	42	39	31	t	204	t	t
<i>pinene</i> ₁₆₀₀	32000	31995	14	8	15	t	17	40	loop
<i>pinene</i> ₃₂₀₀	64000	63995	54	42	60	t	66	123	
<i>robot</i> ₈₀₀	7199	4801	4	3	6	12	i	fail	78
<i>robot</i> ₁₆₀₀	14399	9601	10	10	12	48	t	fail	149
<i>rocket</i> ₆₄₀₀	25601	19200	18	6	18	1918	14	fail	1335
<i>rocket</i> ₁₂₈₀₀	51201	38400	36	17	37	t	37	fail	loop
<i>steering</i> ₆₄₀₀	32000	25601	9	7	12	t	t	10	298
<i>steering</i> ₁₂₈₀₀	32000	25601	19	21	32	t	t	31	

- 198 -

Schlussfolgerung bezüglich restringierter Probleme

- ▶ Wesentlich schwerer als unrestringierte Probleme, Auffinden eines zulässigen Punktes nicht garantiert.
- ▶ SQP Methoden effizient und zuverlässig auf Problemen mittlerer Grösse, d.h. mit einigen hundert Variablen und Restriktionen.
- ▶ Bei noch grösseren Problemen scheinen Innere Punkt Methoden derzeit am effektivsten.
- ▶ Anwendung von SQP und Innere Punktmethoden in Praxis verlangt oftmals (zu) hohen Implementierungsaufwand.
- ▶ Lokale Optimierung \equiv Fine Tuning \neq Strukturelle Optimierung .

- 199 -