Inhaltsübersicht für heute:

Inhaltsübersicht für heute:

Strafverfahren (penalty methods)

Idee: Bestrafe die Verletzung von Nebenbedingungen in der Zielfunktion mit einem Strafterm, der mit einem Strafparameter gewichtet wird, und löse nun eine Folge unrestringierter Probleme mit dieser Zielfunktion für wachsenden Strafparameter, bis die Verletzung klein genug ist.

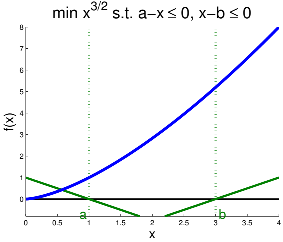
Strafverfahren (penalty methods)

Idee: Bestrafe die Verletzung von Nebenbedingungen in der Zielfunktion mit einem Strafterm, der mit einem Strafparameter gewichtet wird, und löse nun eine Folge unrestringierter Probleme mit dieser Zielfunktion für wachsenden Strafparameter, bis die Verletzung klein genug ist.

Bsp: min $f(x) := x^{\frac{3}{2}}$

s.t. $g_1(x) := a - x \leq 0$

$g_2(x) := x - b \leq 0$



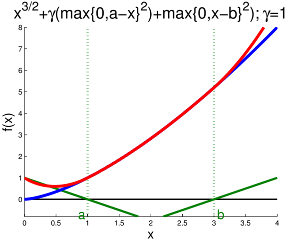
Strafverfahren (penalty methods)

Idee: Bestrafe die Verletzung von Nebenbedingungen in der Zielfunktion mit einem Strafterm, der mit einem Strafparameter gewichtet wird, und löse nun eine Folge unrestringierter Probleme mit dieser Zielfunktion für wachsenden Strafparameter, bis die Verletzung klein genug ist.

Bsp: $\min \quad f(x) := x^{\frac{3}{2}}$

s.t. $\quad g_1(x) := a - x \leq 0$

$\quad g_2(x) := x - b \leq 0$



Unzul. mit Strafparameter $\gamma > 0$ bestrafen:

$f_\gamma(x) := x^{\frac{3}{2}} + \gamma[\max\{0, g_1(x)\}^2 + \max\{0, g_2(x)\}^2]$

Unrestringiertes Problem: $\quad \min_{x \in \mathbb{R}^n} f_\gamma(x)$

Suche Minimum x_γ^*, vergrößere γ, etc.

Strafverfahren (penalty methods)

Idee: Bestrafe die Verletzung von Nebenbedingungen in der Zielfunktion mit einem Strafterm, der mit einem Strafparameter gewichtet wird, und löse nun eine Folge unrestringierter Probleme mit dieser Zielfunktion für wachsenden Strafparameter, bis die Verletzung klein genug ist.

Bsp: $\min \quad f(x) := x^{\frac{3}{2}}$

s.t. $\quad g_1(x) := a - x \le 0$

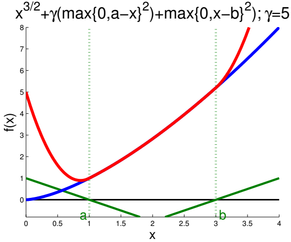
$\quad g_2(x) := x - b \le 0$

Unzul. mit Strafparameter $\gamma > 0$ bestrafen:

$$f_\gamma(x) := x^{\frac{3}{2}} + \gamma[\max\{0, g_1(x)\}^2 + \max\{0, g_2(x)\}^2]$$

Unrestringiertes Problem: $\min_{x \in \mathbb{R}^n} f_\gamma(x)$

Suche Minimum x^*_γ, vergrößere γ, etc.



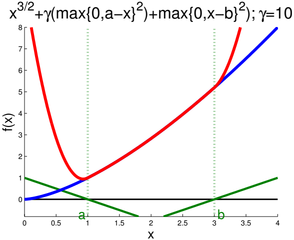
Strafverfahren (penalty methods)

Idee: Bestrafe die Verletzung von Nebenbedingungen in der Zielfunktion mit einem Strafterm, der mit einem Strafparameter gewichtet wird, und löse nun eine Folge unrestringierter Probleme mit dieser Zielfunktion für wachsenden Strafparameter, bis die Verletzung klein genug ist.

Bsp: $\min \quad f(x) := x^{\frac{3}{2}}$

$\text{s.t.} \quad g_1(x) := a - x \leq 0$

$\qquad g_2(x) := x - b \leq 0$



Unzul. mit Strafparameter $\gamma > 0$ bestrafen:

$f_\gamma(x) := x^{\frac{3}{2}} + \gamma[\max\{0, g_1(x)\}^2 + \max\{0, g_2(x)\}^2]$

Unrestringiertes Problem: $\min_{x \in \mathbb{R}^n} f_\gamma(x)$

Suche Minimum x^*_γ, vergrößere γ, etc.

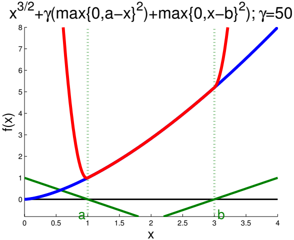
Strafverfahren (penalty methods)

Idee: Bestrafe die Verletzung von Nebenbedingungen in der Zielfunktion mit einem Strafterm, der mit einem Strafparameter gewichtet wird, und löse nun eine Folge unrestringierter Probleme mit dieser Zielfunktion für wachsenden Strafparameter, bis die Verletzung klein genug ist.

Bsp: min $f(x) := x^{\frac{3}{2}}$

s.t. $g_1(x) := a - x \leq 0$

$g_2(x) := x - b \leq 0$



Unzul. mit Strafparameter $\gamma > 0$ bestrafen:

$f_\gamma(x) := x^{\frac{3}{2}} + \gamma[\max\{0, g_1(x)\}^2 + \max\{0, g_2(x)\}^2]$

Unrestringiertes Problem: $\min_{x \in \mathbb{R}^n} f_\gamma(x)$

Suche Minimum x_γ^*, vergrößere γ, etc.

Strafverfahren (penalty methods)

Idee: Bestrafe die Verletzung von Nebenbedingungen in der Zielfunktion mit einem Strafterm, der mit einem Strafparameter gewichtet wird, und löse nun eine Folge unrestringierter Probleme mit dieser Zielfunktion für wachsenden Strafparameter, bis die Verletzung klein genug ist.

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & h_i(x) = 0 \quad i \in \mathcal{E} \\ & g_i(x) \le 0 \quad i \in \mathcal{I} \\ & x \in \mathbb{R}^n \end{array} \quad \to \quad \min_{x\in\mathbb{R}^n} f_\gamma(x) := f(x) + \gamma \Big[\underbrace{\sum_{i\in\mathcal{E}} \Psi(h_i(x)) + \sum_{i\in\mathcal{I}} \Phi(g_i(x))}_{=:\ \sigma(x)} \Big]$$

$=: \sigma(x)$ ... Straffunktion

$\gamma \ge 0$ ... Strafparameter

mit $\begin{cases} \Psi(y) > 0 & \text{für } y \neq 0, \\ \Psi(y) = 0 & \text{für } y = 0, \end{cases}$ und $\begin{cases} \Phi(y) > 0 & \text{für } y > 0, \\ \Phi(y) = 0 & \text{für } y \le 0. \end{cases}$ [Relax.!]

Strafproblem: $\max_{\gamma \ge 0} \Theta(\gamma)$ mit $\Theta(\gamma) := \inf_{x\in\mathbb{R}^n} f_\gamma(x)$

Strafverfahren (penalty methods)

Idee: Bestrafe die Verletzung von Nebenbedingungen in der Zielfunktion mit einem Strafterm, der mit einem Strafparameter gewichtet wird, und löse nun eine Folge unrestringierter Probleme mit dieser Zielfunktion für wachsenden Strafparameter, bis die Verletzung klein genug ist.

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & h_i(x) = 0 \quad i \in \mathcal{E} \\ & g_i(x) \le 0 \quad i \in \mathcal{I} \\ & x \in \mathbb{R}^n \end{array} \quad \rightarrow \quad \min_{x \in \mathbb{R}^n} f_\gamma(x) := f(x) + \gamma \Big[\underbrace{\sum_{i \in \mathcal{E}} \Psi(h_i(x)) + \sum_{i \in \mathcal{I}} \Phi(g_i(x))}_{=: \sigma(x)} \Big]$$

$=: \sigma(x)$ ... Straffunktion

$\gamma \ge 0$ ... Strafparameter

mit $\begin{cases} \Psi(y) > 0 & \text{für } y \neq 0, \\ \Psi(y) = 0 & \text{für } y = 0, \end{cases}$ und $\begin{cases} \Phi(y) > 0 & \text{für } y > 0, \\ \Phi(y) = 0 & \text{für } y \le 0. \end{cases}$ [Relax.!]

Strafproblem: $\max_{\gamma \ge 0} \Theta(\gamma)$ mit $\Theta(\gamma) := \inf_{x \in \mathbb{R}^n} f_\gamma(x)$

Satz (Korrektheit von Strafverfahren)

Ist das Originalproblem zulässig und gibt es für jedes $\gamma \ge 0$ ein globales Optimum x_γ, für das $\Theta(\gamma)$ angenommen wird, und ist $\{x_\gamma : \gamma \ge 0\}$ beschränkt und abgeschlossen, dann sind alle Häufungspunkte von $\{x_\gamma\}$ globale Optimallösungen des Originalproblems.

Exakte Strafverfahren und l_1-Straffunktion

Gibt es für jede lokale Optimallösung x^* des Originalproblems ein endliches $\gamma_{x^*} > 0$, sodass x^* auch lokales Optimum von $f_\gamma(\cdot)$ für alle $\gamma \geq \gamma_{x^*}$ ist, heißt das Strafverfahren und die Straffunktion exakt.

Exakte Strafverfahren und l_1-Straffunktion

Gibt es für jede lokale Optimallösung x^* des Originalproblems ein endliches $\gamma_{x^*} > 0$, sodass x^* auch lokales Optimum von $f_\gamma(\cdot)$ für alle $\gamma \geq \gamma_{x^*}$ ist, heißt das Strafverfahren und die Straffunktion exakt.

Bsp: Die l_1-Straffunktion ist definiert durch

$$\Psi(y) := |y| \text{ und } \Phi(y) := \max\{0, y\}$$

und ist exakt für konvexe Probleme:

Satz (Exaktheit der l_1-Straffunktion für konvexe Probleme)

Sei (x^*, μ^*, λ^*) ein KKT-Punkt eines konvexen Problems (P), dann ist für $\gamma \geq \max(\{|\mu_i^*| : i \in \mathcal{E}\} \cup \{\lambda_i^* : i \in \mathcal{I}\})$ der Punkt x^* auch globales Optimum von $f_\gamma(x) := f(x) + \gamma\big[\sum_{i\in\mathcal{E}} |h_i(x)| + \sum_{i\in\mathcal{I}} \max\{0, g_i(x)\}\big]$.

Exakte Strafverfahren und l_1-Straffunktion

Gibt es für jede lokale Optimallösung x^* des Originalproblems ein endliches $\gamma_{x^*} > 0$, sodass x^* auch lokales Optimum von $f_\gamma(\cdot)$ für alle $\gamma \geq \gamma_{x^*}$ ist, heißt das Strafverfahren und die Straffunktion exakt.

Bsp: Die l_1-Straffunktion ist definiert durch

$$\Psi(y) := |y| \text{ und } \Phi(y) := \max\{0, y\}$$

und ist exakt für konvexe Probleme:

Satz (Exaktheit der l_1-Straffunktion für konvexe Probleme)

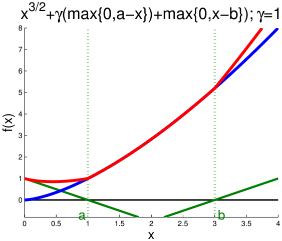
Sei (x^*, μ^*, λ^*) ein KKT-Punkt eines konvexen Problems (P), dann ist für $\gamma \geq \max(\{|\mu_i^*| : i \in \mathcal{E}\} \cup \{\lambda_i^* : i \in \mathcal{I}\})$ der Punkt x^* auch globales Optimum von $f_\gamma(x) := f(x) + \gamma\left[\sum_{i\in\mathcal{E}} |h_i(x)| + \sum_{i\in\mathcal{I}} \max\{0, g_i(x)\}\right]$.

Bsp:
$$\begin{aligned} \min \quad & f(x) := x^{\frac{3}{2}} \\ \text{s.t.} \quad & g_1(x) := a - x \leq 0 \\ & g_2(x) := x - b \leq 0 \end{aligned}$$

$$f_\gamma(x) := x^{\frac{3}{2}} + \gamma[\max\{0, g_1(x)\} + \max\{0, g_2(x)\}]$$

In $x^* = 1$ ist g_1 aktiv mit Lagrangemult. $\lambda_1 = \frac{3}{2}$

[KKT: $0 = \nabla f(1) + \lambda_1 \nabla g_1(1) = \frac{3}{2} - \lambda_1$]



Exakte Strafverfahren und l_1-Straffunktion

Gibt es für jede lokale Optimallösung x^* des Originalproblems ein endliches $\gamma_{x^*} > 0$, sodass x^* auch lokales Optimum von $f_\gamma(\cdot)$ für alle $\gamma \geq \gamma_{x^*}$ ist, heißt das Strafverfahren und die Straffunktion exakt.

Bsp: Die l_1-Straffunktion ist definiert durch

$$\Psi(y) := |y| \text{ und } \Phi(y) := \max\{0, y\}$$

und ist exakt für konvexe Probleme:

Satz (Exaktheit der l_1-Straffunktion für konvexe Probleme)

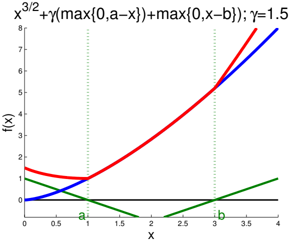
Sei (x^*, μ^*, λ^*) ein KKT-Punkt eines konvexen Problems (P), dann ist für $\gamma \geq \max(\{|\mu_i^*| : i \in \mathcal{E}\} \cup \{\lambda_i^* : i \in \mathcal{I}\})$ der Punkt x^* auch globales Optimum von $f_\gamma(x) := f(x) + \gamma\big[\sum_{i\in\mathcal{E}} |h_i(x)| + \sum_{i\in\mathcal{I}} \max\{0, g_i(x)\}\big]$.

$$\begin{array}{lll} \text{Bsp:} & \min & f(x) := x^{\frac{3}{2}} \\ & \text{s.t.} & g_1(x) := a - x \leq 0 \\ & & g_2(x) := x - b \leq 0 \end{array}$$

$$f_\gamma(x) := x^{\frac{3}{2}} + \gamma[\max\{0, g_1(x)\} + \max\{0, g_2(x)\}]$$

In $x^* = 1$ ist g_1 aktiv mit Lagrangemult. $\lambda_1 = \frac{3}{2}$

[KKT: $0 = \nabla f(1) + \lambda_1 \nabla g_1(1) = \frac{3}{2} - \lambda_1$]



Exakte Strafverfahren und l_1-Straffunktion

Gibt es für jede lokale Optimallösung x^* des Originalproblems ein endliches $\gamma_{x^*} > 0$, sodass x^* auch lokales Optimum von $f_\gamma(\cdot)$ für alle $\gamma \geq \gamma_{x^*}$ ist, heißt das Strafverfahren und die Straffunktion exakt.

Bsp: Die l_1-Straffunktion ist definiert durch

$$\Psi(y) := |y| \text{ und } \Phi(y) := \max\{0, y\}$$

und ist exakt für konvexe Probleme:

Satz (Exaktheit der l_1-Straffunktion für konvexe Probleme)

Sei (x^*, μ^*, λ^*) ein KKT-Punkt eines konvexen Problems (P), dann ist für $\gamma \geq \max(\{|\mu_i^*| : i \in \mathcal{E}\} \cup \{\lambda_i^* : i \in \mathcal{I}\})$ der Punkt x^* auch globales Optimum von $f_\gamma(x) := f(x) + \gamma\left[\sum_{i\in\mathcal{E}} |h_i(x)| + \sum_{i\in\mathcal{I}} \max\{0, g_i(x)\}\right]$.

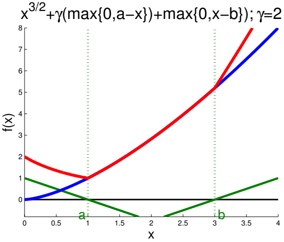
Bsp:

$$\begin{aligned} \min \quad & f(x) := x^{\frac{3}{2}} \\ \text{s.t.} \quad & g_1(x) := a - x \leq 0 \\ & g_2(x) := x - b \leq 0 \end{aligned}$$

$$f_\gamma(x) := x^{\frac{3}{2}} + \gamma[\max\{0, g_1(x)\} + \max\{0, g_2(x)\}]$$

In $x^* = 1$ ist g_1 aktiv mit Lagrangemult. $\lambda_1 = \frac{3}{2}$

[KKT: $0 = \nabla f(1) + \lambda_1 \nabla g_1(1) = \frac{3}{2} - \lambda_1$]



Exakte Strafverfahren und l_1-Straffunktion

Gibt es für jede lokale Optimallösung x^* des Originalproblems ein endliches $\gamma_{x^*} > 0$, sodass x^* auch lokales Optimum von $f_\gamma(\cdot)$ für alle $\gamma \geq \gamma_{x^*}$ ist, heißt das Strafverfahren und die Straffunktion exakt.

Bsp: Die l_1-Straffunktion ist definiert durch

$$\Psi(y) := |y| \text{ und } \Phi(y) := \max\{0, y\}$$

und ist exakt für konvexe Probleme:

Satz (Exaktheit der l_1-Straffunktion für konvexe Probleme)

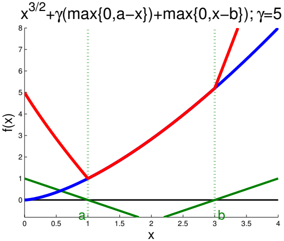
Sei (x^*, μ^*, λ^*) ein KKT-Punkt eines konvexen Problems (P), dann ist für $\gamma \geq \max(\{|\mu_i^*| : i \in \mathcal{E}\} \cup \{\lambda_i^* : i \in \mathcal{I}\})$ der Punkt x^* auch globales Optimum von $f_\gamma(x) := f(x) + \gamma\left[\sum_{i\in\mathcal{E}} |h_i(x)| + \sum_{i\in\mathcal{I}} \max\{0, g_i(x)\}\right]$.

Bsp:
$$\begin{array}{lll} \min & f(x) := x^{\frac{3}{2}} \\ \text{s.t.} & g_1(x) := a - x \leq 0 \\ & g_2(x) := x - b \leq 0 \end{array}$$

$$f_\gamma(x) := x^{\frac{3}{2}} + \gamma[\max\{0, g_1(x)\} + \max\{0, g_2(x)\}]$$

In $x^* = 1$ ist g_1 aktiv mit Lagrangemult. $\lambda_1 = \frac{3}{2}$

$[\text{KKT: } 0 = \nabla f(1) + \lambda_1 \nabla g_1(1) = \frac{3}{2} - \lambda_1]$



Exakte Strafverfahren und l_1-Straffunktion

Gibt es für jede lokale Optimallösung x^* des Originalproblems ein endliches $\gamma_{x^*} > 0$, sodass x^* auch lokales Optimum von $f_\gamma(\cdot)$ für alle $\gamma \geq \gamma_{x^*}$ ist, heißt das Strafverfahren und die Straffunktion exakt.

Bsp: Die l_1-Straffunktion ist definiert durch

$$\Psi(y) := |y| \text{ und } \Phi(y) := \max\{0, y\}$$

und ist exakt für konvexe Probleme:

Satz (Exaktheit der l_1-Straffunktion für konvexe Probleme)

Sei (x^*, μ^*, λ^*) ein KKT-Punkt eines konvexen Problems (P), dann ist für $\gamma \geq \max(\{|\mu_i^*| : i \in \mathcal{E}\} \cup \{\lambda_i^* : i \in \mathcal{I}\})$ der Punkt x^* auch globales Optimum von $f_\gamma(x) := f(x) + \gamma\left[\sum_{i\in\mathcal{E}} |h_i(x)| + \sum_{i\in\mathcal{I}} \max\{0, g_i(x)\}\right]$.

Was hat γ mit den Lagrangemultiplikatoren zu tun?
Um x^* bestraft $\mathcal{L}(x, \mu^*, \lambda^*) = f(x) + \sum \mu_i^* h_i(x) + \sum \lambda_i^* g_i(x)$ unzulässige Richtungen, in die sich f verbessert, mit $\mu_i^* h_i(x) \geq 0$ bzw. $\lambda_i^* g_i(x) \geq 0$ gerade so, dass der Anstieg $\nabla_x \mathcal{L}(x^*, \mu^*, \lambda^*) = 0$ ist. Jeder größere Anstieg γ macht es nur noch unattraktiver.

Nachteil der l_1-Straffuntkion: f_γ ist nicht differenzierbar!

Die quadratische Straffunktion (quadratic penalty)

Verwende $\Psi(y) := \frac{1}{2}y^2$ und $\Phi(y) := \frac{1}{2}\max\{0, y\}^2$, also

$$f_\gamma(x) = f(x) + \frac{\gamma}{2}\left[\sum_{i\in\mathcal{E}} h_i(x)^2 + \sum_{i\in\mathcal{I}} \max\{0, g_i(x)\}^2\right].$$

- f_γ ist stetig diffbar, aber für $\mathcal{I} \neq \emptyset$ i.A. nicht zweimal stetig diffbar.
- Falls $\mathcal{I} = \emptyset$, kann jedes glatte unrestringierte Verfahren zur Bestimmung eines x_γ^* verwendet werden.
- Gerade für Gleichungsnebenbedingungen führt die quadratische Straffunktion zu bananenförmigen Teilproblemen $\longrightarrow$ ungünstige Konvergenzeigenschaften.
- Für große γ wird die zweite Ableitung in unzulässige Richtungen sehr groß $\longrightarrow$ numerische Probleme.
- Die quadratische Straffunktion ist i.A. nicht exakt, selbst für sehr großes γ bleibt x_γ^* unzulässig (s. Anfangsbeispiel).

Inhaltsübersicht für heute:

Augmentierte-Lagrange-Verfahren (nur Gleichungen)

Idee: addiere die quadratische Straffunktion zur Lagrange-Funktion,

$$\mathcal{L}_\gamma(x,\mu) := f(x) + \sum_{i\in\mathcal{E}} \mu_i h_i(x) + \frac{\gamma}{2}\sum_{i\in\mathcal{E}} h_i(x)^2 \quad = \mathcal{L}(x,\mu) + \frac{\gamma}{2}\|h(x)\|^2,$$

dann bleibt die Funktion diffbar und ist für Multiplikator μ^* exakt.

Augmentierte-Lagrange-Verfahren (nur Gleichungen)

Idee: addiere die quadratische Straffunktion zur Lagrange-Funktion,

$$\mathcal{L}_\gamma(x,\mu) := f(x) + \sum_{i\in\mathcal{E}} \mu_i h_i(x) + \frac{\gamma}{2}\sum_{i\in\mathcal{E}} h_i(x)^2 \quad = \mathcal{L}(x,\mu) + \frac{\gamma}{2}\|h(x)\|^2,$$

dann bleibt die Funktion diffbar und ist für Multiplikator μ^* exakt.

Erfüllt x^* die hinreichenden Opt.-bed. mit Multiplikator μ^*, dann gilt

$$\nabla_x\mathcal{L}(x^*,\mu^*) = 0 \text{ und } d^T\nabla_{xx}\mathcal{L}(x^*,\mu^*)d > 0 \quad \forall d \in T_P(x^*) [\Leftrightarrow J_h(x^*)d = 0].$$

Augmentierte-Lagrange-Verfahren (nur Gleichungen)

Idee: addiere die quadratische Straffunktion zur Lagrange-Funktion,

$$\mathcal{L}_\gamma(x,\mu) := f(x) + \sum_{i\in\mathcal{E}} \mu_i h_i(x) + \frac{\gamma}{2}\sum_{i\in\mathcal{E}} h_i(x)^2 \quad = \mathcal{L}(x,\mu) + \frac{\gamma}{2}\|h(x)\|^2,$$

dann bleibt die Funktion diffbar und ist für Multiplikator μ^* exakt.

Erfüllt x^* die hinreichenden Opt.-bed. mit Multiplikator μ^*, dann gilt

$$\nabla_x\mathcal{L}(x^*,\mu^*) = 0 \text{ und } d^T\nabla_{xx}\mathcal{L}(x^*,\mu^*)d > 0 \quad \forall d \in T_P(x^*)[\Leftrightarrow J_h(x^*)d = 0].$$

In allen Richtungen orthogonal dazu $[d = J_h(x^*)^T u \text{ für } u \in \mathbb{R}^{\mathcal{E}}]$ sorgt

wegen $\nabla_{xx}\mathcal{L}_\gamma(x,\mu) = \nabla_{xx}\mathcal{L}(x,\mu) + \gamma\sum_{i\in\mathcal{E}} \underbrace{h_i(x)}_{=0 \text{ in } x^*} \nabla^2 h_i(x) + \gamma J_h(x)^T J_h(x)$

der Summand $J_h(x^*)^T J_h(x^*)$ für γ groß genug dafür, dass insgesamt

$$\nabla_x\mathcal{L}_\gamma(x^*,\mu^*) = 0 \text{ und } \nabla_{xx}\mathcal{L}_\gamma(x^*,\mu^*) \succ 0.$$

$\Rightarrow$ Für γ groß genug erfüllt x^* die hinr. freien Opt.-Bed. für $\mathcal{L}_\gamma(\cdot,\mu^*)$.

Augmentierte-Lagrange-Verfahren (nur Gleichungen)

Idee: addiere die quadratische Straffunktion zur Lagrange-Funktion,

$$\mathcal{L}_\gamma(x,\mu) := f(x) + \sum_{i\in\mathcal{E}} \mu_i h_i(x) + \frac{\gamma}{2}\sum_{i\in\mathcal{E}} h_i(x)^2 \quad = \mathcal{L}(x,\mu) + \frac{\gamma}{2}\|h(x)\|^2,$$

dann bleibt die Funktion diffbar und ist für Multiplikator μ^* exakt.

Erfüllt x^* die hinreichenden Opt.-bed. mit Multiplikator μ^*, dann gilt

$$\nabla_x\mathcal{L}(x^*,\mu^*) = 0 \text{ und } d^T\nabla_{xx}\mathcal{L}(x^*,\mu^*)d > 0 \quad \forall d \in T_P(x^*)[\Leftrightarrow J_h(x^*)d = 0].$$

In allen Richtungen orthogonal dazu $[d = J_h(x^*)^T u$ für $u \in \mathbb{R}^{\mathcal{E}}]$ sorgt

$$\text{wegen } \nabla_{xx}\mathcal{L}_\gamma(x,\mu) = \nabla_{xx}\mathcal{L}(x,\mu) + \gamma\sum_{i\in\mathcal{E}} \underbrace{h_i(x)}_{=0 \text{ in } x^*} \nabla^2 h_i(x) + \gamma J_h(x)^T J_h(x)$$

der Summand $J_h(x^*)^T J_h(x^*)$ für γ groß genug dafür, dass insgesamt

$$\nabla_x\mathcal{L}_\gamma(x^*,\mu^*) = 0 \text{ und } \nabla_{xx}\mathcal{L}_\gamma(x^*,\mu^*) \succ 0.$$

$\Rightarrow$ Für γ groß genug erfüllt x^* die hinr. freien Opt.-Bed. für $\mathcal{L}_\gamma(\cdot,\mu^*)$.

Geometrisch: $\mathcal{L}(\cdot,\mu^*)$ ist wegen $\nabla_x\mathcal{L}_\gamma(x^*,\mu^*) = 0$ in alle Richtungen um x^* flach und in zulässige Richtungen lokal konvex. Der Strafterm $\frac{\gamma}{2}\|h(x)\|^2$ macht die Funktion in den unzulässigen Richtungen lokal konvex.

$$\mathcal{L}_\gamma(x,\mu) := f(x) + \sum_{i\in\mathcal{E}} \mu_i h_i(x) + \frac{\gamma}{2}\sum_{i\in\mathcal{E}} h_i(x)^2 \quad = \mathcal{L}(x,\mu) + \frac{\gamma}{2}\|h(x)\|^2$$

Satz (Augmentierte-Lagrange-Verfahren bei Gleichungen)

Erfüllt x^* die hinreichenden Optimalitätsbedingungen für ein restringiertes Problem ohne Ungleichungen mit Multiplikator μ^*, dann gibt es ein $\underline{\gamma} > 0$, sodass x^* für jedes $\gamma > \underline{\gamma}$ ein lokales Minimum von $\min_{x\in\mathbb{R}^n} \mathcal{L}_\gamma(x,\mu^*)$ *ist, das die hinreichenden freien Optimalitätsbedingungen erfüllt.*

$$\mathcal{L}_\gamma(x,\mu) := f(x) + \sum_{i\in\mathcal{E}} \mu_i h_i(x) + \frac{\gamma}{2}\sum_{i\in\mathcal{E}} h_i(x)^2 \quad = \mathcal{L}(x,\mu) + \frac{\gamma}{2}\|h(x)\|^2$$

Satz (Augmentierte-Lagrange-Verfahren bei Gleichungen)

Erfüllt x^* die hinreichenden Optimalitätsbedingungen für ein restringiertes Problem ohne Ungleichungen mit Multiplikator μ^*, dann gibt es ein $\underline{\gamma} > 0$, sodass x^* für jedes $\gamma > \underline{\gamma}$ ein lokales Minimum von $\min_{x\in\mathbb{R}^n} \mathcal{L}_\gamma(x,\mu^*)$ ist, das die hinreichenden freien Optimalitätsbedingungen erfüllt.

Ist μ^* bekannt $\rightarrow$ lokale freie Optimierung. Aber μ^* ist unbekannt!
Wie wählt man μ im Verfahren? (iterativ, Schritt $k \;\rightarrow\; k+1$)

$$\mathcal{L}_\gamma(x,\mu) := f(x) + \sum_{i\in\mathcal{E}} \mu_i h_i(x) + \frac{\gamma}{2}\sum_{i\in\mathcal{E}} h_i(x)^2 \quad = \mathcal{L}(x,\mu) + \frac{\gamma}{2}\|h(x)\|^2$$

Satz (Augmentierte-Lagrange-Verfahren bei Gleichungen)

Erfüllt x^* die hinreichenden Optimalitätsbedingungen für ein restringiertes Problem ohne Ungleichungen mit Multiplikator μ^*, dann gibt es ein $\underline{\gamma} > 0$, sodass x^* für jedes $\gamma > \underline{\gamma}$ ein lokales Minimum von $\min_{x\in\mathbb{R}^n} \mathcal{L}_\gamma(x,\mu^*)$ ist, das die hinreichenden freien Optimalitätsbedingungen erfüllt.

Ist μ^* bekannt $\rightarrow$ lokale freie Optimierung. Aber μ^* ist unbekannt!
Wie wählt man μ im Verfahren? (iterativ, Schritt $k \rightarrow k+1$)
μ^* und x^* erfüllen $0 = \nabla_x \mathcal{L}(x^*,\mu^*) = \nabla f(x^*) + \sum_{i\in\mathcal{E}} \mu_i^* \nabla h_i(x^*)$.

$$\mathcal{L}_\gamma(x,\mu) := f(x) + \sum_{i\in\mathcal{E}} \mu_i h_i(x) + \frac{\gamma}{2}\sum_{i\in\mathcal{E}} h_i(x)^2 \quad = \mathcal{L}(x,\mu) + \frac{\gamma}{2}\|h(x)\|^2$$

Satz (Augmentierte-Lagrange-Verfahren bei Gleichungen)

Erfüllt x^* die hinreichenden Optimalitätsbedingungen für ein restringiertes Problem ohne Ungleichungen mit Multiplikator μ^*, dann gibt es ein $\underline{\gamma} > 0$, sodass x^* für jedes $\gamma > \underline{\gamma}$ ein lokales Minimum von $\min_{x\in\mathbb{R}^n} \mathcal{L}_\gamma(x,\mu^*)$ ist, das die hinreichenden freien Optimalitätsbedingungen erfüllt.

Ist μ^* bekannt $\rightarrow$ lokale freie Optimierung. Aber μ^* ist unbekannt!
Wie wählt man μ im Verfahren? (iterativ, Schritt $k \rightarrow k+1$)
μ^* und x^* erfüllen $0 = \nabla_x \mathcal{L}(x^*,\mu^*) = \nabla f(x^*) + \sum_{i\in\mathcal{E}} \mu_i^* \nabla h_i(x^*)$.

Sei $x^{(k)}$ die berechnete Lösung zu $\min_{x\in\mathbb{R}^n} \mathcal{L}_{\gamma_k}(x,\mu^{(k)})$, dann ist

$$0 \approx \nabla_x \mathcal{L}_{\gamma_k}(x^{(k)},\mu^{(k)}) = \nabla f(x^{(k)}) + \sum_{i\in\mathcal{E}} \underbrace{[\mu_i^{(k)} + \gamma_k h_i(x^{(k)})]}_{\approx \mu^*} \nabla h_i(x^{(k)})$$

$$\mathcal{L}_\gamma(x,\mu) := f(x) + \sum_{i\in\mathcal{E}} \mu_i h_i(x) + \frac{\gamma}{2}\sum_{i\in\mathcal{E}} h_i(x)^2 \quad = \mathcal{L}(x,\mu) + \frac{\gamma}{2}\|h(x)\|^2$$

Satz (Augmentierte-Lagrange-Verfahren bei Gleichungen)

Erfüllt x^* die hinreichenden Optimalitätsbedingungen für ein restringiertes Problem ohne Ungleichungen mit Multiplikator μ^*, dann gibt es ein $\underline{\gamma} > 0$, sodass x^* für jedes $\gamma > \underline{\gamma}$ ein lokales Minimum von $\min_{x\in\mathbb{R}^n} \mathcal{L}_\gamma(x,\mu^*)$ ist, das die hinreichenden freien Optimalitätsbedingungen erfüllt.

Ist μ^* bekannt $\to$ lokale freie Optimierung. Aber μ^* ist unbekannt!
Wie wählt man μ im Verfahren? (iterativ, Schritt $k \to k+1$)
μ^* und x^* erfüllen $0 = \nabla_x \mathcal{L}(x^*,\mu^*) = \nabla f(x^*) + \sum_{i\in\mathcal{E}} \mu_i^* \nabla h_i(x^*)$.

Sei $x^{(k)}$ die berechnete Lösung zu $\min_{x\in\mathbb{R}^n} \mathcal{L}_{\gamma_k}(x,\mu^{(k)})$, dann ist

$$0 \approx \nabla_x \mathcal{L}_{\gamma_k}(x^{(k)},\mu^{(k)}) = \nabla f(x^{(k)}) + \sum_{i\in\mathcal{E}} \underbrace{[\mu_i^{(k)} + \gamma_k h_i(x^{(k)})]}_{\approx\mu^*} \nabla h_i(x^{(k)})$$

$$\to \quad \mu_i^{(k+1)} := \mu_i^{(k)} + \gamma_k h_i(x^{(k)}) \qquad (i\in\mathcal{E})$$

Alg. Schema für Augm.-Lagr.-Verf. mit Gleichungen

0. Wähle Startpunkt $\bar{x}^{(0)}$, $\mu^{(0)}$, $\gamma_0 > 0$, Genauigkeit $\varepsilon_0 > 0$, $k := 0$
1. Bestimme Näherungslösung $x^{(k)}$ zu $\quad \min_{x \in \mathbb{R}^n} \mathcal{L}_{\gamma_k}(x, \mu^{(k)}) \quad$ mit Startpunkt $\bar{x}^{(k)}$, so dass $\|\nabla_x \mathcal{L}_{\gamma_k}(x^{(k)}, \mu^{(k)})\| \leq \varepsilon_k$.
2. Ist $\|h(x^{(k)})\|$ und $\|\nabla_x \mathcal{L}_(x^{(k)}, \mu^{(k)})\|$ klein genug, STOP. (u.U. Neustart, falls kein Fortschritt Richtung Zulässigkeit, etc.)
3. Aktualisiere die Multiplikatoren: $\mu_i^{(k+1)} := \mu_i^{(k)} + \gamma_k h_i(x^{(k)}) \qquad (i \in \mathcal{E})$
4. Wähle einen neuen Strafparameter $\gamma_{k+1} \in [\gamma_k, \infty)$
5. Wähle die nächste Genauigkeit $\varepsilon_{k+1} \in (0, \varepsilon_k]$
6. Wähle den nächsten Startpunkt, meist $\bar{x}^{(k+1)} := x^{(k)}$
7. $k \leftarrow k + 1$, GOTO 1.

Alg. Schema für Augm.-Lagr.-Verf. mit Gleichungen

0. Wähle Startpunkt $\bar{x}^{(0)}$, $\mu^{(0)}$, $\gamma_0 > 0$, Genauigkeit $\varepsilon_0 > 0$, $k := 0$
1. Bestimme Näherungslösung $x^{(k)}$ zu $\min_{x\in\mathbb{R}^n} \mathcal{L}_{\gamma_k}(x, \mu^{(k)})$ mit Startpunkt $\bar{x}^{(k)}$, so dass $\|\nabla_x \mathcal{L}_{\gamma_k}(x^{(k)}, \mu^{(k)})\| \le \varepsilon_k$.
2. Ist $\|h(x^{(k)})\|$ und $\|\nabla_x \mathcal{L}_{(}x^{(k)}, \mu^{(k)})\|$ klein genug, STOP. (u.U. Neustart, falls kein Fortschritt Richtung Zulässigkeit, etc.)
3. Aktualisiere die Multiplikatoren: $\mu_i^{(k+1)} := \mu_i^{(k)} + \gamma_k h_i(x^{(k)}) \qquad (i \in \mathcal{E})$
4. Wähle einen neuen Strafparameter $\gamma_{k+1} \in [\gamma_k, \infty)$
5. Wähle die nächste Genauigkeit $\varepsilon_{k+1} \in (0, \varepsilon_k]$
6. Wähle den nächsten Startpunkt, meist $\bar{x}^{(k+1)} := x^{(k)}$
7. $k \leftarrow k + 1$, GOTO 1.

Ungleichungen:
werden z.B. in LANCELOT mit Schlupfvariablen in Glgen umgewandelt:

$$g_i(x) + s_i = 0,\ s_i \ge 0 \quad (i \in \mathcal{I})$$

In Schritt 1 wird das Problem dann mit Vorzeichenbedingungen gelöst,

$$\min_{x\in\mathbb{R}^n, s\in\mathbb{R}_+^{\mathcal{I}}} \mathcal{L}_{\gamma_k}(x, s, \tilde{\mu}^{(k)})$$

(z.B. über ein quadratisches Modell mit linearen Unglgen, s. später).

Inhaltsübersicht für heute:

Strafverfahren

Augmentierte-Lagrange-Verfahren

Barriere-Verfahren

Barriere-Verfahren

Idee: Verhindere das Verlassen des zulässigen Bereiches durch einen Barriere-Term in der Zielfunktion.

$\rightarrow$ vor allem für Ungleichungen gut geeignet! Im Folgenden $\mathcal{E} = \emptyset$.

Barriere-Verfahren

Idee: Verhindere das Verlassen des zulässigen Bereiches durch einen Barriere-Term in der Zielfunktion.

$\rightarrow$ vor allem für Ungleichungen gut geeignet! Im Folgenden $\mathcal{E} = \emptyset$.

Bsp:

Barriere-Verfahren

Idee: Verhindere das Verlassen des zulässigen Bereiches durch einen Barriere-Term in der Zielfunktion.

$\rightarrow$ vor allem für Ungleichungen gut geeignet!   Im Folgenden $\mathcal{E} = \emptyset$.

Bsp: Innere-Punkte-Verfahren für Lineare Optimierung über Kegeln

Barriere-Verfahren

Idee: Verhindere das Verlassen des zulässigen Bereiches durch einen Barriere-Term in der Zielfunktion.

$\rightarrow$ vor allem für Ungleichungen gut geeignet! Im Folgenden $\mathcal{E} = \emptyset$.

Bsp: Innere-Punkte-Verfahren für Lineare Optimierung über Kegeln

---

Das zulässige Innere $\mathring{\mathcal{X}} := \{x \in \mathbb{R}^n : g_i(x) < 0, i \in \mathcal{I}\}$ sei nicht leer, $\mathring{\mathcal{X}} \neq \emptyset$.
Barriere-Funktionen erfüllen folgende Eigenschaften:

- Für $x \notin \mathring{\mathcal{X}}$ haben sie den Wert ∞.
- Innerhalb von $\mathring{\mathcal{X}}$ sind sie glatt.
- Geht eine Folge $x^{(k)} \in \mathring{\mathcal{X}}$ gegen den Rand von $\mathring{\mathcal{X}}$, geht der Wert gegen ∞.

Barriere-Verfahren

Idee: Verhindere das Verlassen des zulässigen Bereiches durch einen Barriere-Term in der Zielfunktion.

$\rightarrow$ vor allem für Ungleichungen gut geeignet! Im Folgenden $\mathcal{E} = \emptyset$.

Bsp: Innere-Punkte-Verfahren für Lineare Optimierung über Kegeln

---

Das zulässige Innere $\mathring{\mathcal{X}} := \{x \in \mathbb{R}^n : g_i(x) < 0, i \in \mathcal{I}\}$ sei nicht leer, $\mathring{\mathcal{X}} \neq \emptyset$.
Barriere-Funktionen erfüllen folgende Eigenschaften:

- Für $x \notin \mathring{\mathcal{X}}$ haben sie den Wert ∞.
- Innerhalb von $\mathring{\mathcal{X}}$ sind sie glatt.
- Geht eine Folge $x^{(k)} \in \mathring{\mathcal{X}}$ gegen den Rand von $\mathring{\mathcal{X}}$, geht der Wert gegen ∞.

Beispiele für $y < 0$: $-\frac{1}{y}$, $-\log(-y)$

Barriere-Verfahren

Idee: Verhindere das Verlassen des zulässigen Bereiches durch einen Barriere-Term in der Zielfunktion.

$\rightarrow$ vor allem für Ungleichungen gut geeignet! Im Folgenden $\mathcal{E} = \emptyset$.

Bsp: Innere-Punkte-Verfahren für Lineare Optimierung über Kegeln

---

Das zulässige Innere $\mathring{\mathcal{X}} := \{x \in \mathbb{R}^n : g_i(x) < 0, i \in \mathcal{I}\}$ sei nicht leer, $\mathring{\mathcal{X}} \neq \emptyset$.
Barriere-Funktionen erfüllen folgende Eigenschaften:

- Für $x \notin \mathring{\mathcal{X}}$ haben sie den Wert ∞.
- Innerhalb von $\mathring{\mathcal{X}}$ sind sie glatt.
- Geht eine Folge $x^{(k)} \in \mathring{\mathcal{X}}$ gegen den Rand von $\mathring{\mathcal{X}}$, geht der Wert gegen ∞.

Beispiele für $y < 0$: $-\frac{1}{y}$, $-\log(-y)$

Der Einfluss der Barriere-Funktion wird durch einen Barriereparameter β kontrolliert und man sucht ein lokales Minimum für

$$\min_{x \in \mathbb{R}^n} f_\beta(x) := f(x) - \beta \sum_{i \in \mathcal{I}} \log(-g_i(x))$$

ausgehend von einem Startpunkt im Inneren.

Alg. Schema für Barriere-Verf. (nur Unglgen)

0. Wähle Startpunkt $\bar{x}^{(0)} \in \mathring{\mathcal{X}}$, $\beta_0 > 0$, Genauigkeit $\varepsilon_0 > 0$, $k := 0$
1. Bestimme Näherungslösung $x^{(k)}$ zu $\min_{x\in\mathbb{R}^n} f_{\beta_k}(x)$ mit Startpunkt $\bar{x}^{(k)}$, so dass $\|\nabla_x f_{\beta_k}(x^{(k)})\| \leq \varepsilon_k$.
2. Sind die KKT-Bedingungen näherungsweise erfüllt, STOP. (teste ob Lösungen nach unendlich gehen, etc.)
3. Wähle einen neuen Barriereparameter $\beta_{k+1} \in (0, \beta_k)$
4. Wähle die nächste Genauigkeit $\varepsilon_{k+1} \in (0, \varepsilon_k]$
5. Wähle den nächsten Startpunkt, meist $\bar{x}^{(k+1)} := x^{(k)}$
6. $k \leftarrow k + 1$, GOTO 1.

Alg. Schema für Barriere-Verf. (nur Unglgen)

0. Wähle Startpunkt $\bar{x}^{(0)} \in \mathring{\mathcal{X}}$, $\beta_0 > 0$, Genauigkeit $\varepsilon_0 > 0$, $k := 0$
1. Bestimme Näherungslösung $x^{(k)}$ zu $\min_{x \in \mathbb{R}^n} f_{\beta_k}(x)$ mit Startpunkt $\bar{x}^{(k)}$, so dass $\|\nabla_x f_{\beta_k}(x^{(k)})\| \leq \varepsilon_k$.
2. Sind die KKT-Bedingungen näherungsweise erfüllt, STOP. (teste ob Lösungen nach unendlich gehen, etc.)
3. Wähle einen neuen Barriereparameter $\beta_{k+1} \in (0, \beta_k)$
4. Wähle die nächste Genauigkeit $\varepsilon_{k+1} \in (0, \varepsilon_k]$
5. Wähle den nächsten Startpunkt, meist $\bar{x}^{(k+1)} := x^{(k)}$
6. $k \leftarrow k+1$, GOTO 1.

KKT-Bedingungen? Sei $x(\beta)$ lokale OL zu $\min_{x \in \mathbb{R}^n} f_\beta(x)$, dann ist

$$0 = \nabla f_\beta(x(\beta)) \quad = \quad \nabla f(x(\beta)) + \sum_{i \in \mathcal{I}} \underbrace{\frac{\beta}{-g_i(x(\beta))}}_{=:\lambda_i(\beta)\ \geq 0} \nabla g_i(x(\beta))$$

Alg. Schema für Barriere-Verf. (nur Unglgen)

0. Wähle Startpunkt $\bar{x}^{(0)} \in \mathring{\mathcal{X}}$, $\beta_0 > 0$, Genauigkeit $\varepsilon_0 > 0$, $k := 0$
1. Bestimme Näherungslösung $x^{(k)}$ zu $\min_{x\in\mathbb{R}^n} f_{\beta_k}(x)$ mit Startpunkt $\bar{x}^{(k)}$, so dass $\|\nabla_x f_{\beta_k}(x^{(k)})\| \le \varepsilon_k$.
2. Sind die KKT-Bedingungen näherungsweise erfüllt, STOP. (teste ob Lösungen nach unendlich gehen, etc.)
3. Wähle einen neuen Barriereparameter $\beta_{k+1} \in (0, \beta_k)$
4. Wähle die nächste Genauigkeit $\varepsilon_{k+1} \in (0, \varepsilon_k]$
5. Wähle den nächsten Startpunkt, meist $\bar{x}^{(k+1)} := x^{(k)}$
6. $k \leftarrow k + 1$, GOTO 1.

KKT-Bedingungen? Sei $x(\beta)$ lokale OL zu $\min_{x\in\mathbb{R}^n} f_\beta(x)$, dann ist

$$0 = \nabla f_\beta(x(\beta)) \quad = \quad \nabla f(x(\beta)) + \sum_{i\in\mathcal{I}} \underbrace{\frac{\beta}{-g_i(x(\beta))}}_{=:\lambda_i(\beta)\ \ge 0} \nabla g_i(x(\beta))$$

also ist

$$\begin{aligned} \nabla f(x(\beta)) + \sum_{i\in\mathcal{I}} \lambda_i(\beta)\nabla g_i(x(\beta)) &= 0 \\ g_i(x(\beta)) &\le 0 \quad (i \in \mathcal{I}) \\ \lambda_i(\beta) &\ge 0 \quad (i \in \mathcal{I}) \\ -\lambda_i(\beta) g_i(x(\beta)) &= \beta \quad (i \in \mathcal{I}) \text{ [pert. Kompl.]} \end{aligned}$$

Alg. Schema für Barriere-Verf. (nur Unglgen)

0. Wähle Startpunkt $\bar{x}^{(0)} \in \overset{\circ}{\mathcal{X}}$, $\beta_0 > 0$, Genauigkeit $\varepsilon_0 > 0$, $k := 0$
1. Bestimme Näherungslösung $x^{(k)}$ zu $\min_{x\in\mathbb{R}^n} f_{\beta_k}(x)$ mit Startpunkt $\bar{x}^{(k)}$, so dass $\|\nabla_x f_{\beta_k}(x^{(k)})\| \le \varepsilon_k$.
2. Sind die KKT-Bedingungen näherungsweise erfüllt, STOP. (teste ob Lösungen nach unendlich gehen, etc.)
3. Wähle einen neuen Barriereparameter $\beta_{k+1} \in (0, \beta_k)$
4. Wähle die nächste Genauigkeit $\varepsilon_{k+1} \in (0, \varepsilon_k]$
5. Wähle den nächsten Startpunkt, meist $\bar{x}^{(k+1)} := x^{(k)}$
6. $k \leftarrow k+1$, GOTO 1.

KKT-Bedingungen? Sei $x(\beta)$ lokale OL zu $\min_{x\in\mathbb{R}^n} f_\beta(x)$, dann ist

$$0 = \nabla f_\beta(x(\beta)) \quad = \quad \nabla f(x(\beta)) + \sum_{i\in\mathcal{I}} \underbrace{\frac{\beta}{-g_i(x(\beta))}}_{=:\lambda_i(\beta)\ \geq 0} \nabla g_i(x(\beta))$$

also ist

$$\begin{aligned} \nabla f(x(\beta)) + \sum_{i\in\mathcal{I}} \lambda_i(\beta)\nabla g_i(x(\beta)) &= 0 \\ g_i(x(\beta)) &\le 0 \quad (i\in\mathcal{I}) \\ \lambda_i(\beta) &\ge 0 \quad (i\in\mathcal{I}) \\ -\lambda_i(\beta) g_i(x(\beta)) &= \beta \quad (i\in\mathcal{I}) \ \text{[pert. Kompl.]} \end{aligned}$$

Für $\beta \to 0$ wird bei Konvergenz ein KKT-Punkt gefunden!

Satz (Fiacco und Mc Cormick 1968)

Sei $\mathring{\mathcal{X}} \neq \emptyset$ und x^* ein lokales Minimum, in dem (LICQ) und die hinreichenden Optimalitätsbed. mit streng komplementären Multiplikatoren λ^* erfüllt sind. Dann gilt:

(i) Für $\bar{\beta}$ klein genug gibt es eine diffbare Funktion $x(\beta) : (0, \bar{\beta}) \to \mathbb{R}^n$ mit $x(\beta) \overset{\beta \to 0}{\longrightarrow} x^*$ und $x(\beta)$ ist lokales Minimum von f_β. [zent. Pfad]

(ii) Für $x(\beta)$ aus (i) konvergieren die Schätzer der Lagrange-Multiplikatoren $\lambda_i(\beta) = \frac{\beta}{-g_i(x(\beta))} \overset{\beta \to 0}{\longrightarrow} \lambda_i^*$ für $i \in \mathcal{I}$.

(iii) Für β klein genug erfüllt die Hessematrix $\nabla_x^2 f_\beta(x(\beta)) \succ 0$.

Satz (Fiacco und Mc Cormick 1968)

Sei $\mathring{\mathcal{X}} \neq \emptyset$ und x^* ein lokales Minimum, in dem (LICQ) und die hinreichenden Optimalitätsbed. mit streng komplementären Multiplikatoren λ^* erfüllt sind. Dann gilt:

(i) Für $\bar{\beta}$ klein genug gibt es eine diffbare Funktion $x(\beta) : (0, \bar{\beta}) \to \mathbb{R}^n$ mit $x(\beta) \overset{\beta \to 0}{\to} x^*$ und $x(\beta)$ ist lokales Minimum von f_β. [zent. Pfad]

(ii) Für $x(\beta)$ aus (i) konvergieren die Schätzer der Lagrange-Multiplikatoren $\lambda_i(\beta) = \frac{\beta}{-g_i(x(\beta))} \overset{\beta \to 0}{\to} \lambda_i^*$ für $i \in \mathcal{I}$.

(iii) Für β klein genug erfüllt die Hessematrix $\nabla_x^2 f_\beta(x(\beta)) \succ 0$.

Bemerkungen:

- Die Glattheit von $x(\beta)$ kann zur Beschleunigung durch Prädiktor-Korrektor-Verfahren genutzt werden.

Satz (Fiacco und Mc Cormick 1968)

Sei $\mathring{\mathcal{X}} \neq \emptyset$ und x^* ein lokales Minimum, in dem (LICQ) und die hinreichenden Optimalitätsbed. mit streng komplementären Multiplikatoren λ^* erfüllt sind. Dann gilt:

(i) Für $\bar{\beta}$ klein genug gibt es eine diffbare Funktion $x(\beta) : (0, \bar{\beta}) \to \mathbb{R}^n$ mit $x(\beta) \overset{\beta \to 0}{\to} x^*$ und $x(\beta)$ ist lokales Minimum von f_β. [zent. Pfad]

(ii) Für $x(\beta)$ aus (i) konvergieren die Schätzer der Lagrange-Multiplikatoren $\lambda_i(\beta) = \frac{\beta}{-g_i(x(\beta))} \overset{\beta \to 0}{\to} \lambda_i^*$ für $i \in \mathcal{I}$.

(iii) Für β klein genug erfüllt die Hessematrix $\nabla_x^2 f_\beta(x(\beta)) \succ 0$.

Bemerkungen:

- Die Glattheit von $x(\beta)$ kann zur Beschleunigung durch Prädiktor-Korrektor-Verfahren genutzt werden.
- $\nabla_x^2 f_\beta$ wird auch schlecht skaliert, aber das ist theoretisch begründbar numerisch meist noch akzeptabel.

Satz (Fiacco und Mc Cormick 1968)

Sei $\mathring{\mathcal{X}} \neq \emptyset$ und x^* ein lokales Minimum, in dem (LICQ) und die hinreichenden Optimalitätsbed. mit streng komplementären Multiplikatoren λ^* erfüllt sind. Dann gilt:

(i) Für $\bar{\beta}$ klein genug gibt es eine diffbare Funktion $x(\beta) : (0, \bar{\beta}) \to \mathbb{R}^n$ mit $x(\beta) \overset{\beta \to 0}{\to} x^*$ und $x(\beta)$ ist lokales Minimum von f_β. [zent. Pfad]

(ii) Für $x(\beta)$ aus (i) konvergieren die Schätzer der Lagrange-Multiplikatoren $\lambda_i(\beta) = \frac{\beta}{-g_i(x(\beta))} \overset{\beta \to 0}{\to} \lambda_i^*$ für $i \in \mathcal{I}$.

(iii) Für β klein genug erfüllt die Hessematrix $\nabla_x^2 f_\beta(x(\beta)) \succ 0$.

Bemerkungen:

- Die Glattheit von $x(\beta)$ kann zur Beschleunigung durch Prädiktor-Korrektor-Verfahren genutzt werden.
- $\nabla_x^2 f_\beta$ wird auch schlecht skaliert, aber das ist theoretisch begründbar numerisch meist noch akzeptabel.
- Gleichungsnebenbedingungen werden meist durch einen Strafterm eingebunden.