# Statische Prozessoptimierung/ Prozessoptimierung 1

## Kapitel 7:
## Nichtlineare Optimierung beschränkter Probleme

**Prof. Dr.-Ing. habil Pu Li**

**Fachgebiet Prozessoptimierung**

# Optimierungsprobleme mit Nebenbedingungen:



## Allgemeine Darstellung:

$$\min_{\mathbf{x},\mathbf{u}} f(\mathbf{x},\mathbf{u})$$

mit
$$\mathbf{g}(\mathbf{x},\mathbf{u}) = \mathbf{0}$$
$$\mathbf{h}(\mathbf{x},\mathbf{u}) \geq \mathbf{0}$$
$$\mathbf{x}_{min} \leq \mathbf{x} \leq \mathbf{x}_{max}$$
$$\mathbf{u}_{min} \leq \mathbf{u} \leq \mathbf{u}_{max}$$

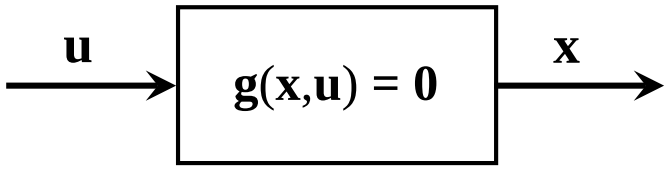
wobei
- $\mathbf{x}$: *n* Zustandsvariablen (abhängige Variablen)
- $\mathbf{u}$: *m* Steuervariablen (unabhängige Variablen)
- $\mathbf{g}$: *n* Modellgleichungen
- $\mathbf{h}$: *l* Beschränkungen

## Simultane Betrachtung:

Das Problem wird in dieser Form betrachtet und direkt mit einem Lösungsverfahren gelöst. Bei der Lösung sind alle Variablen ($\mathbf{x}$ und $\mathbf{u}$) zu optimieren.

## Sequentielle Betrachtung:

Die Zustandsvariablen hängen von den Steuervariablen ab.




Man kann durch

1. **Umformung und Ersetzung**
2. **Simulation**

die Zustandsvariablen $\mathbf{x}$ und die Gleichungen $\mathbf{g}$ eliminieren. Also gibt es $\mathbf{x} = \varphi(\mathbf{u})$ und nun ist das Problem

$$\min_{\mathbf{u}} \; f(\mathbf{x}(\mathbf{u}), \mathbf{u})$$

$$\text{mit} \quad \mathbf{h}(\mathbf{x}(\mathbf{u}), \mathbf{u}) \geq \mathbf{0}$$

$$\mathbf{x}_{min} \leq \mathbf{x}(\mathbf{u}) \leq \mathbf{x}_{max}$$

$$\mathbf{u}_{min} \leq \mathbf{u} \leq \mathbf{u}_{max}$$

Da nur die Steuervariablen auftauchen, stellt es ein kleines Optimierungs-problem dar. Es gibt keine Gleichungsnebenbedingungen.

Wenn das originale Problem keine Ungleichungsnebenbedingungen hat, ist das Problem

$$\min_{\mathbf{u}} f(\mathbf{x}(\mathbf{u}), \mathbf{u}) = f(\mathbf{u}),$$

das mit dem Newton-, dem Gradienten- oder dem Quasi-Newton-Verfahren gelöst werden kann.

**Bei Umformung/Ersetzung** muss man aufpassen, manchmal werden Gleichungen nicht erfüllt!

**Beispiel:**

$$\min_{\mathbf{x}} f(x_1, x_2) = x_1^2 + x_2^2$$

$$(x_1 - 1)^3 - x_2^2 = 0$$

Die Lösung des Problems ist $x_1^* = 1,\ x_2^* = 0.$

Aber wenn man die Gleichung umformt, d. h.

$$x_2^2 = (x_1 - 1)^3$$

und diese in die Zielfunktion einsetzt, erhält man

$$\min_{x_1} f(x_1) = x_1^2 + (x_1 - 1)^3.$$

# Nichtlineare Optimierung: zwei Variablen mit einer Nebenbedingung

$$\min_{\mathbf{x}} \; f(x_1, x_2)$$

mit $\quad g(x_1, x_2) = 0$

Bei der Lösung $(x_1^*, x_2^*)$ gibt es:

$$df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 = 0$$

$$dg = \frac{\partial g}{\partial x_1} dx_1 + \frac{\partial g}{\partial x_2} dx_2 = 0$$

weil $dx_1, dx_2 \neq 0$, ergibt sich

$$\frac{\partial f}{\partial x_1}\frac{\partial g}{\partial x_2} - \frac{\partial f}{\partial x_2}\frac{\partial g}{\partial x_1} = 0$$

d. h.

$$\frac{\left.\frac{\partial f}{\partial x_1}\right|_*}{\left.\frac{\partial g}{\partial x_1}\right|_*} = \frac{\left.\frac{\partial f}{\partial x_2}\right|_*}{\left.\frac{\partial g}{\partial x_2}\right|_*} = \lambda$$

dann
$$\frac{\partial f}{\partial x_1} = \lambda \frac{\partial g}{\partial x_1} \qquad \frac{\partial f}{\partial x_2} = \lambda \frac{\partial g}{\partial x_2}$$

Es bedeutet

$$\frac{\partial f}{\partial x_1} - \lambda \frac{\partial g}{\partial x_1} = \frac{\partial}{\partial x_1}\left[f(x_1, x_2) - \lambda\, g(x_1, x_2)\right] = 0$$
$$\frac{\partial f}{\partial x_2} - \lambda \frac{\partial g}{\partial x_2} = \frac{\partial}{\partial x_2}\left[f(x_1, x_2) - \lambda\, g(x_1, x_2)\right] = 0$$

die Defínition der Lagrange-Funktion:

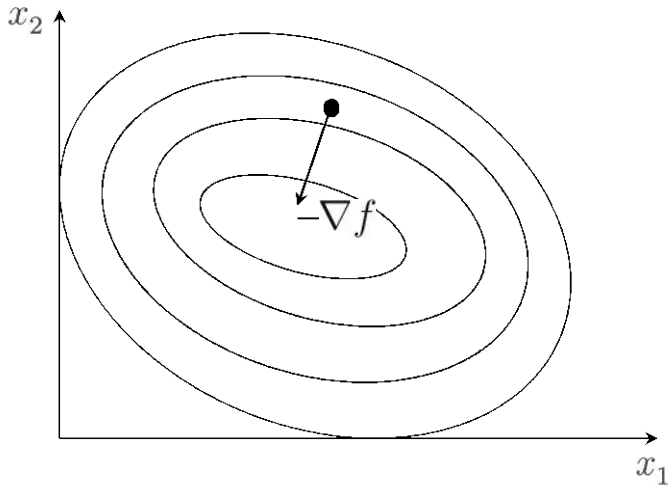
$$L(x_1, x_2, \lambda) = f(x_1, x_2) - \lambda\, g(x_1, x_2)$$

$\lambda$ heißt Lagrange-Multiplikator.

Zusätzlich:
$$\begin{bmatrix} \dfrac{\partial f}{\partial x_1} \\ \dfrac{\partial f}{\partial x_2} \end{bmatrix} - \lambda \begin{bmatrix} \dfrac{\partial g}{\partial x_1} \\ \dfrac{\partial g}{\partial x_2} \end{bmatrix} = \nabla f - \lambda \nabla g = \mathbf{0}$$

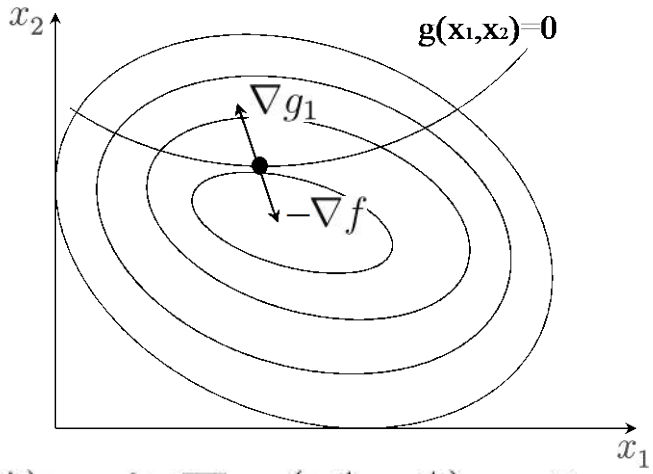
Bei der Lösung sind die zwei Vektoren parallel und entgegengesetzt gerichtet.

## NLP ohne Nebenbedingungen:




Es ergibt sich: $\nabla f(x_1^*, x_2^*) = 0, \ (\Delta \mathbf{x})^T \nabla f^2(x_1^*, x_2^*)(\Delta \mathbf{x}) \geq 0$

## NLP mit einer Gleichungsnebenbedingung:

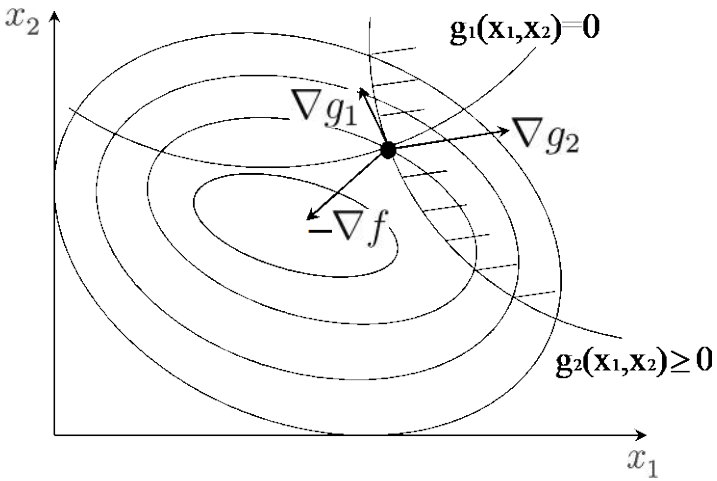



Dann

$$\nabla f(x_1^*, x_2^*) - \lambda_1 \nabla g_1(x_1^*, x_2^*) = 0$$

$$(\Delta \mathbf{x})^T \left[\nabla^2 f(x_1^*, x_2^*) - \lambda_1 \nabla^2 g_1(x_1^*, x_2^*)\right](\Delta \mathbf{x}) \geq 0$$

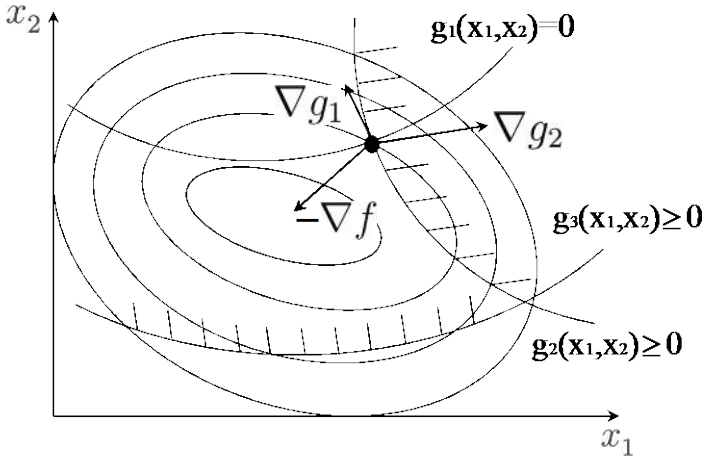
## NLP mit einer Gleichungs- und einer Ungleichungsnebenbedingung:




$g_1$ und $g_2$ sind aktiv, $\lambda_2 > 0$

$$\nabla f(x_1^*, x_2^*) - \lambda_1 \nabla g_1(x_1^*, x_2^*) - \lambda_2 \nabla g_2(x_1^*, x_2^*) = 0$$

## NLP mit einer Gleichungs- und zwei Ungleichungsnebenbedingungen:




$g_1$ und $g_2$ sind aktiv, $\lambda_2 > 0,\ \lambda_3 = 0$

$$\nabla f(x_1^*, x_2^*) - \lambda_1 \nabla g_1(x_1^*, x_2^*) - \lambda_2 \nabla g_2(x_1^*, x_2^*) - \lambda_3 \nabla g_3(x_1^*, x_2^*) = 0$$

# Minimierung mit einer Gleichungsnebenbedingung

$$\min_{\mathbf{x}} \; x_1^2 + x_2^2$$

mit $x_1 + x_2 = 1$

$$f(x_1, x_2) = x_1^2 + x_2^2, \quad g(x_1, x_2) = x_1 + x_2 - 1 = 0$$

dann $\nabla f(x_1, x_2) = \begin{bmatrix} 2x_1 & 2x_2 \end{bmatrix}^T, \quad \nabla g(x_1, x_2) = \begin{bmatrix} 1 & 1 \end{bmatrix}^T$

Der Lösungspunkt:

$$x_1^* = x_2^* = \frac{1}{2}, \quad \nabla g^* = \begin{bmatrix} 1 & 1 \end{bmatrix}^T, \quad \nabla f^* = \begin{bmatrix} 1 & 1 \end{bmatrix}^T$$

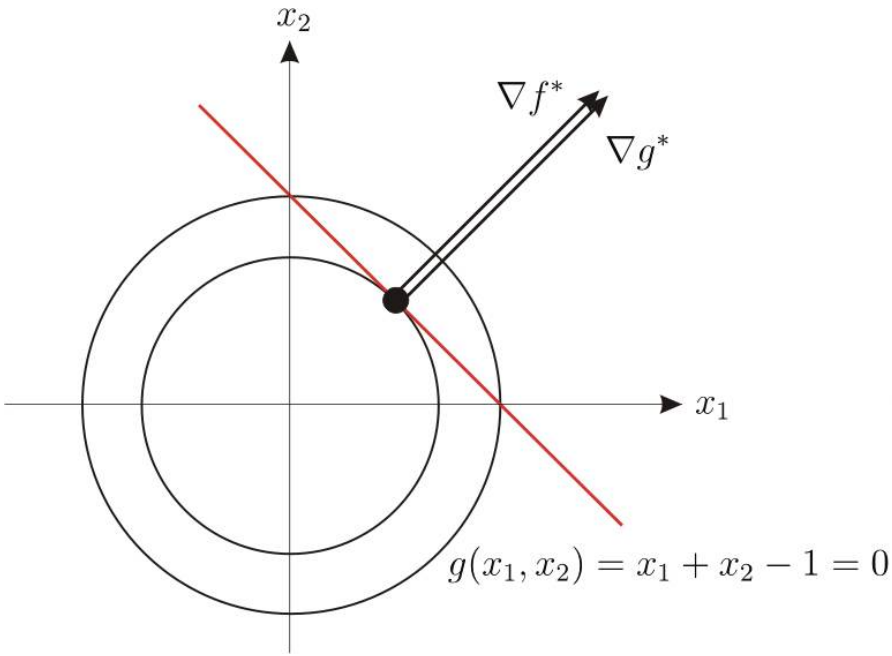
D. h. $\nabla f^* = \nabla g^*$ also $\lambda^* = 1 > 0$

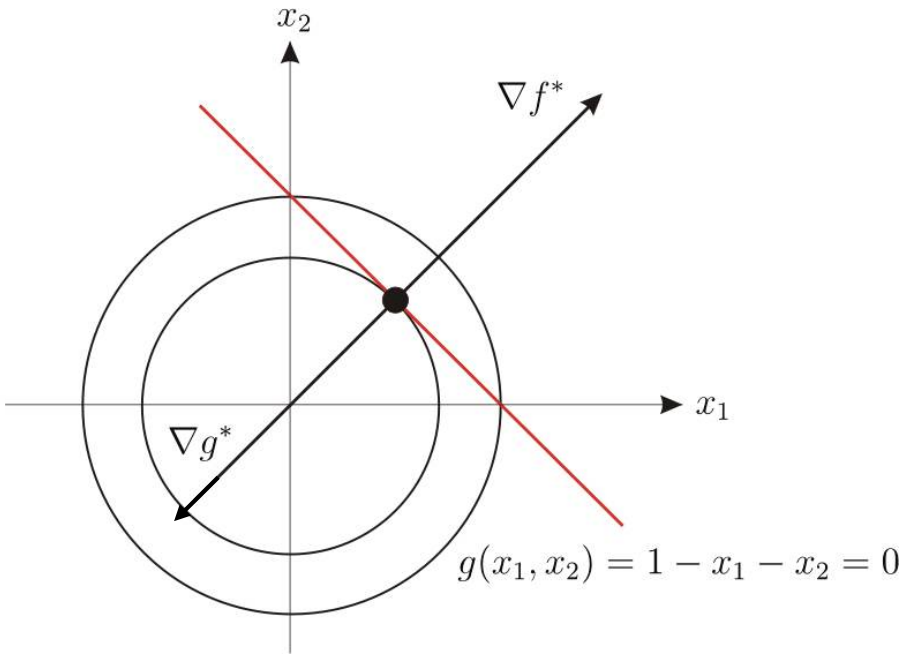
Aber man kann auch schreiben $g(x_1, x_2) = 1 - x_1 - x_2 = 0$,

damit $\nabla g(x_1, x_2) = \begin{bmatrix} -1 & -1 \end{bmatrix}^T$. Der Lösungspunkt bleibt

$x_1^* = x_2^* = \frac{1}{2}$, daher $\nabla f^* = -\nabla g^*, \; \lambda^* = -1 < 0$.

# Grafische Darstellung:



## Problemdarstellung:

Die Zielfunktion und die Nebenbedingungen sind nichtlineare Funktionen des Variablenvektors.

$$\min_{\mathbf{x}} f(\mathbf{x})$$

mit

$$g_j(\mathbf{x}) = 0 \qquad j = 1, \cdots, m \qquad \text{(I)}$$

$$g_j(\mathbf{x}) \geq 0 \qquad j = m+1, \cdots, m+r \qquad \text{(II)}$$

$$x_{i\,min} \leq x_i \leq x_{i\,max} \qquad i = 1, \cdots, N$$

## Karush-Kuhn-Tucker-Bedingungen:

An der Lösung $\mathbf{x} = \mathbf{x}^*$ müssen die folgenden Bedingungen erfüllt werden:

- $\nabla L(\mathbf{x}^*, \boldsymbol{\lambda}^*) = 0$ (I)
- $g_j(\mathbf{x}^*) = 0 \qquad j = 1, \cdots, m$ (II)
- $g_j(\mathbf{x}^*) \geq 0 \qquad j = m+1, \cdots, m+r$ (II)
- $\lambda_j^* g_j(\mathbf{x}^*) = 0, \quad \lambda_j^* \geq 0 \quad j = m+1, \cdots, m+r$
- $\nabla g_j(\mathbf{x}^*)$ voneinander unabhängig $j \in A$
- $(\Delta\mathbf{x})^T [\nabla^2 L(\mathbf{x}^*, \boldsymbol{\lambda}^*)](\Delta\mathbf{x}) \geq 0$

# Quadratische Programmierung mit linearen Gleichungsnebenbedingungen

## Das QPE-Problem

$$\min_{\mathbf{x}} \left\{ f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{Q}\,\mathbf{x} + \mathbf{c}^T\mathbf{x} \right\}$$

mit $\quad \mathbf{g}(\mathbf{x}) = \mathbf{A}\,\mathbf{x} - \mathbf{b} = \mathbf{0}$

wobei $\mathbf{Q} = \mathbf{Q}^T \in R^{n \times n}$ eine positiv definite Hesse-Matrix ist und:

$$\begin{aligned} \mathbf{x}, \mathbf{c} &\in R^n \\ \mathbf{A} &\in R^{m \times n} \\ \mathbf{g}(\mathbf{x}), \mathbf{b} &\in R^m \\ Rang\ (\mathbf{A}) &= m \end{aligned}$$

$n$ : Variablen

$m$: Gleichungen

## Analytische Lösung

Lagrange-Funktion: $$L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) - \boldsymbol{\lambda}^T(\mathbf{A}\mathbf{x} - \mathbf{b})$$

KKT-Bedingungen: $$\nabla L = \mathbf{Q}\mathbf{x} + \mathbf{c} - \mathbf{A}^T\boldsymbol{\lambda} = \mathbf{0}$$
$$\mathbf{A}\mathbf{x} - \mathbf{b} = \mathbf{0}$$

Hieraus folgt $$\mathbf{Q}\mathbf{x} - \mathbf{A}^T\boldsymbol{\lambda} = -\mathbf{c}$$
$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

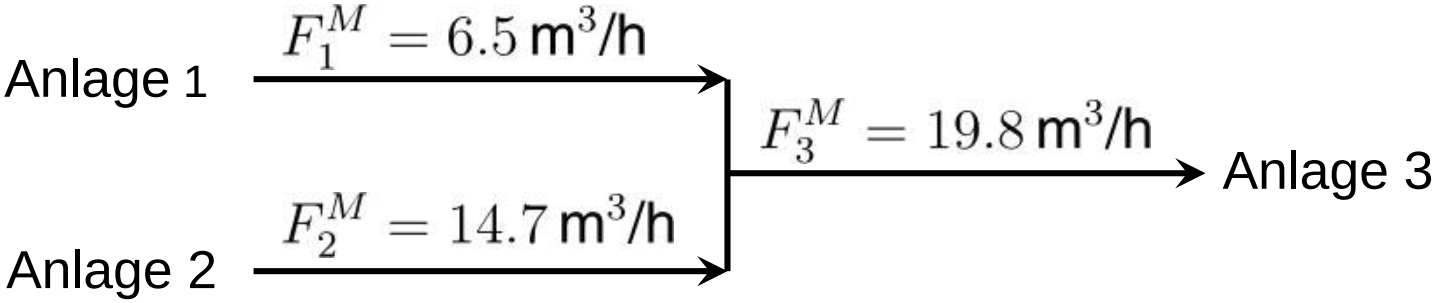
d. h. $$\begin{bmatrix} \mathbf{Q} & \mathbf{A}^T \\ \mathbf{A} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ -\boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} -\mathbf{c} \\ \mathbf{b} \end{bmatrix}$$

Wenn $\mathbf{Q}$ positiv definit ist und $Rang\ (\mathbf{A}) = m$

$$\begin{bmatrix} \mathbf{x}^* \\ -\boldsymbol{\lambda}^* \end{bmatrix} = \begin{bmatrix} \mathbf{Q} & \mathbf{A}^T \\ \mathbf{A} & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} -\mathbf{c} \\ \mathbf{b} \end{bmatrix}$$

$$\mathbf{x}^* = -\mathbf{Q}^{-1}\mathbf{c} + \mathbf{Q}^{-1}\mathbf{A}^T(\mathbf{A}\,\mathbf{Q}^{-1}\mathbf{A}^T)^{-1}(\mathbf{A}\,\mathbf{Q}^{-1}\mathbf{c} + \mathbf{b})$$
$$\boldsymbol{\lambda}^* = (\mathbf{A}\,\mathbf{Q}^{-1}\mathbf{A}^T)^{-1}(\mathbf{A}\,\mathbf{Q}^{-1}\mathbf{c} + \mathbf{b})$$

# Beispiel: Messdatenvalidierung




**Probleme:**
1) Die Daten sind nicht konsistent.
2) Es führt zu Streit.

**Lösung:** Optimierung (Messwertvalidierung)

$$\min_{F_1,F_2,F_3} \left(\frac{F_1 - F_1^M}{\sigma_1}\right)^2 + \left(\frac{F_2 - F_2^M}{\sigma_2}\right)^2 + \left(\frac{F_3 - F_3^M}{\sigma_3}\right)^2$$

mit $F_1 + F_2 = F_3$

wobei die Standardabweichung der Messgeräte:

$$\sigma_1 = 0.1\,\mathrm{m^3/h}$$
$$\sigma_2 = 0.2\,\mathrm{m^3/h}$$
$$\sigma_3 = 0.3\,\mathrm{m^3/h}$$

## Die Lagrange-Funktion:

$$L(F_1, F_2, F_3, \lambda) = \left(\frac{F_1 - F_1^M}{\sigma_1}\right)^2 + \left(\frac{F_2 - F_2^M}{\sigma_2}\right)^2 + \left(\frac{F_3 - F_3^M}{\sigma_3}\right)^2 - \lambda(F_1 + F_2 - F_3)$$

Nach der Optimalitätsbedingung:

$$\frac{\partial L}{\partial F_1} = 2\left(\frac{F_1 - F_1^M}{\sigma_1^2}\right) - \lambda = 0$$

$$\frac{\partial L}{\partial F_2} = 2\left(\frac{F_2 - F_2^M}{\sigma_2^2}\right) - \lambda = 0$$

$$\frac{\partial L}{\partial F_3} = 2\left(\frac{F_3 - F_3^M}{\sigma_3^2}\right) - \lambda = 0$$

$$F_1 + F_2 - F_3 = 0$$

## Die validierte Messung:

$$F_1 = F_1^M + \frac{\sigma_1^2\left(F_3^M - F_1^M - F_2^M\right)}{\sigma_1^2 + \sigma_2^2 + \sigma_3^2} = 6.4\,\mathrm{m^3/h}$$

$$F_2 = F_2^M + \frac{\sigma_2^2\left(F_3^M - F_1^M - F_2^M\right)}{\sigma_1^2 + \sigma_2^2 + \sigma_3^2} = 14.3\,\mathrm{m^3/h}$$

$$F_3 = F_3^M + \frac{\sigma_3^2\left(F_1^M + F_2^M - F_3^M\right)}{\sigma_1^2 + \sigma_2^2 + \sigma_3^2} = 20.7\,\mathrm{m^3/h}$$

## Erwartungswert dieser Variablen:

$$\mu_1^V = E(F_1) = E\left[F_1^M + \frac{\sigma_1^2\left(F_3^M - F_1^M - F_2^M\right)}{\sigma_1^2 + \sigma_2^2 + \sigma_3^2}\right] = \mu_1 + \frac{\sigma_1^2\left(\mu_3 - \mu_1 - \mu_2\right)}{\sigma_1^2 + \sigma_2^2 + \sigma_3^2}$$

## Varianz der Variablen:

$$\left(\sigma_1^V\right)^2 = E\left[\left(F_1 - \mu_1^V\right)^2\right] = E\left[F_1^M - \mu_1 + \frac{\sigma_1^2\left\{\left(F_3^M - \mu_3\right) - \left(F_1^M - \mu_2\right) - \left(F_2^M - \mu_1\right)\right\}}{\sigma_1^2 + \sigma_2^2 + \sigma_3^2}\right]^2$$

## Daraus ergibt sich:

$$\left(\sigma_1^V\right)^2 = \frac{\sigma_1^2\left(\sigma_2^2 + \sigma_3^2\right)}{\sigma_1^2 + \sigma_2^2 + \sigma_3^2}$$

## dann ist die validierte Standardabweichung:

$$\sigma_1^V = \sigma_1\sqrt{\frac{\sigma_2^2 + \sigma_3^2}{\sigma_1^2 + \sigma_2^2 + \sigma_3^2}} = 0.096$$

$$\sigma_2^V = \sigma_2\sqrt{\frac{\sigma_1^2 + \sigma_3^2}{\sigma_1^2 + \sigma_2^2 + \sigma_3^2}} = 0.169$$

$$\sigma_3^V = \sigma_3\sqrt{\frac{\sigma_1^2 + \sigma_2^2}{\sigma_1^2 + \sigma_2^2 + \sigma_3^2}} = 0.179$$

# Quadratische Programmierung mit linearen Gleichungs- und Ungleichungsnebenbedingungen

Das QP-Problem:

$$\min_{\mathbf{x}} \; f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} + \mathbf{c}^T\mathbf{x}$$

$$\text{mit} \quad g_i(\mathbf{x}) = \mathbf{a}_i^T\mathbf{x} - b_i = 0 \qquad i \in I$$

$$g_i(\mathbf{x}) = \mathbf{a}_i^T\mathbf{x} - b_i \geq 0 \qquad i \in II$$

Lagrange-Funktion:

$$L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) - \boldsymbol{\lambda}^T(\mathbf{A}\mathbf{x} - \mathbf{b})$$

$$= f(\mathbf{x}) - \sum_{i=1}^{m+r} \lambda_i(\mathbf{a}_i^T\mathbf{x} - b_i)$$

Da das Problem auch Ungleichungsnebenbedingungen besitzt, kann es nicht analytisch gelöst werden. Es wird mit dem sog. „Aktive-Restriktionen-Verfahren“ (active set method) numerisch iterativ gelöst.

Man fängt mit $\mathbf{x}^0$ an, berechnet die Änderung $\mathbf{d}$ und dann

$$\mathbf{x}^1 = \mathbf{x}^0 + \mathbf{d}$$

**Definition 1:** Ein Punkt $x^k$ ist zulässig, wenn er für die beiden Nebenbedingungen gültig ist.

**Definition 2:** Die aktive Menge der Restriktionen in der Iteration $k$ ist

$$A^k = \{j \in \quad II, \quad g_j(\mathbf{x}^k) = \mathbf{a}_j^T \mathbf{x}^k - b_j = 0\}$$

Taylor-Entwicklung von $x^k$ mit der Änderung $\mathbf{d}$

$$f(\mathbf{x}^k + \mathbf{d}) = f(\mathbf{x}^k) + (\mathbf{Q}\mathbf{x}^k + \mathbf{c})^T \mathbf{d} + \frac{1}{2}\mathbf{d}^T \mathbf{Q}\mathbf{d}$$

$$g_i(\mathbf{x}^k + \mathbf{d}) = g_i(\mathbf{x}^k) + \mathbf{a}_i^T \mathbf{d} \qquad i \in I \cup II$$

d. h. $$L_{xx}^2 = \mathbf{Q}, \quad \nabla f(\mathbf{x}^k) = \mathbf{Q}\,\mathbf{x}^k + \mathbf{c}, \quad \nabla g_i = \mathbf{a}_i$$

Das neue Optimierungsproblem:

$$\min_{\mathbf{d}} \ \frac{1}{2}\mathbf{d}^T \mathbf{Q}\mathbf{d} + (\mathbf{Q}\mathbf{x}^k + \mathbf{c})^T \mathbf{d}$$

mit $\mathbf{a}_i^T \mathbf{d} = 0 \qquad i \in A^k$

$g_i(\mathbf{x}^k) + \mathbf{a}_i^T \mathbf{d} > 0 \qquad i \notin A^k$

## Die Zielfunktion:

$$f(\mathbf{d}+\mathbf{x}^k) = \frac{1}{2}(\mathbf{d}+\mathbf{x}^k)^T\mathbf{Q}(\mathbf{d}+\mathbf{x}^k) + \mathbf{c}^T(\mathbf{d}+\mathbf{x}^k)$$
$$= \frac{1}{2}\mathbf{d}^T\mathbf{Q}\mathbf{d} + \frac{1}{2}\mathbf{d}^T\mathbf{Q}\mathbf{x}^k + \frac{1}{2}(\mathbf{x}^k)^T\mathbf{Q}\mathbf{d} + \frac{1}{2}(\mathbf{x}^k)^T\mathbf{Q}\mathbf{x}^k + \mathbf{c}^T\mathbf{d} + \mathbf{c}^T\mathbf{x}^k$$
$$= \frac{1}{2}\mathbf{d}^T\mathbf{Q}\mathbf{d} + (\mathbf{Q}\mathbf{x}^k + \mathbf{c})^T\mathbf{d} + \frac{1}{2}(\mathbf{x}^k)^T\mathbf{Q}\mathbf{x}^k + \mathbf{c}^T\mathbf{x}^k$$

## Die Nebenbedingungen:

$$g_i(\mathbf{d}+\mathbf{x}^k) = \mathbf{a}_i^T(\mathbf{d}+\mathbf{x}^k) - b_i$$
$$= \mathbf{a}_i^T\mathbf{d} + \mathbf{a}_i^T\mathbf{x}^\mathbf{k} - b_i = \mathbf{a}_i^T\mathbf{d} + g_i(\mathbf{x}^k) > 0$$

## Das Optimierungsproblem:

$$\min_{\mathbf{d}} \ \frac{1}{2}\mathbf{d}^T\mathbf{Q}\mathbf{d} + (\mathbf{Q}\mathbf{x}^k + \mathbf{c})^T\mathbf{d}$$

mit $\mathbf{a}_i^T\mathbf{d} = 0 \qquad i \in A^k$

$\mathbf{a}_i^T\mathbf{d} + g_i(\mathbf{x}^k) \geq 0 \qquad i \notin A^k$

# Der Algorithmus (aktive set method)



1) Einen zulässigen Punkt $\mathbf{x}^0$ auswählen. Die entsprechende Menge der aktiven Restriktionen identifizieren, $k = 0$
2) Das folgende QPE-Problem lösen

$$\min_{\mathbf{d}} \ \frac{1}{2}\mathbf{d}^T\mathbf{Q}\mathbf{d} + (\mathbf{Q}\mathbf{x}^k + \mathbf{c})^T\mathbf{d}$$

$$\text{mit} \quad \mathbf{a}_i^T\mathbf{d} = 0 \qquad i \in A^k$$

wenn $\mathbf{d} = \mathbf{0}$, GOTO 3)
wenn $\mathbf{d} \neq \mathbf{0}$, GOTO 4)

3) Lagrange-Multiplikatoren prüfen:

$$L(\mathbf{d}, \boldsymbol{\lambda}) = \frac{1}{2}\mathbf{d}^T\mathbf{Q}\mathbf{d} + (\mathbf{Q}\mathbf{x}^k + \mathbf{c})^T\mathbf{d} - \boldsymbol{\lambda}^T(\mathbf{A}_i\mathbf{d})$$

$$\nabla_{\mathbf{d}} L(\mathbf{d}, \boldsymbol{\lambda}) = \mathbf{Q}\mathbf{d} + \mathbf{Q}\mathbf{x}^k + \mathbf{c} - \mathbf{A}_i^T\boldsymbol{\lambda} = \mathbf{0}$$

$$\nabla L(\mathbf{0}, \boldsymbol{\lambda}) = \mathbf{Q}\mathbf{x}^k + \mathbf{c} - \mathbf{A}_i^T\boldsymbol{\lambda} = \mathbf{0}$$

$$\mathbf{Q}\mathbf{x}^k + \mathbf{c} = \sum_i \lambda_i \mathbf{a}_i$$

$$\lambda_j = min\{\lambda_i, i \in A^k\}$$

wenn $\lambda_j \geq 0$, dann $x^* = x^k$, STOP
wenn $\lambda_j < 0$, dann $A^k = A^k - \{j\}$, GOTO 2)

4) Modifikation des berechneten Änderungsschritts

Für die nicht aktiven Nebenbedingungen wird benötigt:

$$g_i(\mathbf{x}^k + \mathbf{d}) = g_i(\mathbf{x}^k) + \mathbf{a}_i^T \mathbf{d} > 0 \qquad i \notin A^k$$

weil $x^k$ zulässig ist, dann $g_i(\mathbf{x}^k) > 0$.
Wenn $\mathbf{a}_i^T \mathbf{d} \geq 0$, braucht $\mathbf{d}$ nicht zu modifizieren.
Wenn $\mathbf{a}_i^T \mathbf{d} < 0$, dann $\mathbf{x}^{k+1} = \mathbf{x}^k + \alpha \mathbf{d} \qquad 0 \leq \alpha \leq 1$

d. h.

$$\mathbf{a}_i^T(\mathbf{x}^k + \alpha \mathbf{d}) - b_i > 0$$

$$\mathbf{a}_i^T \mathbf{x}^k + \alpha\, \mathbf{a}_i^T \mathbf{d} > b_i$$

$$\alpha\, \mathbf{a}_i^T \mathbf{d} > b_i - \mathbf{a}_i^T \mathbf{x}^k$$

$$\alpha < \frac{b_i - \mathbf{a}_i^T \mathbf{x}^k}{\mathbf{a}_i^T \mathbf{d}}$$

wenn $\alpha \geq 1$, dann $\alpha = 1$ und $A^{k+1} = A^k$
wenn $\alpha < 1$, dann

$$\alpha_p = min \left\{ \frac{b_i - \mathbf{a}_i^T \mathbf{x}^k}{\mathbf{a}_i^T \mathbf{d}}, \quad i \notin A^k \right\}$$

Eine neue Ungleichungsnebenbedingung ist aktiv.

$$A^{k+1} = A^k \cup \{p\}$$

5) $k = k + 1$, GOTO 2)

Beispiel:

$$\min_{\mathbf{x}} \; f(\mathbf{x}) = x_1^2 + x_2^2 - 4x_1 - 5x_2 + 2$$

mit $g_1(\mathbf{x}) = -2x_1 - x_2 + 2 \geq 0$ (1)

$g_2(\mathbf{x}) = x_1 \geq 0$ (2)

$g_3(\mathbf{x}) = x_2 \geq 0$ (3)

Standardform der Zielfunktion:

$$f(\mathbf{x}) = \frac{1}{2}\begin{bmatrix} x_1 & x_2 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} -4 & -5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + 2$$

$$\mathbf{Q} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \qquad \mathbf{c}^T = \begin{bmatrix} -4 & -5 \end{bmatrix}$$

Die Nebenbedingungen:

$$\begin{bmatrix} g_1(\mathbf{x}) \\ g_2(\mathbf{x}) \\ g_3(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} -2 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \mathbf{a}_3^T \end{bmatrix}\mathbf{x} + \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \geq \mathbf{0}$$

Gradienten der Zielfunktion:

$$\nabla f(\mathbf{x}) = \begin{bmatrix} \dfrac{\partial f}{\partial x_1} \\ \dfrac{\partial f}{\partial x_1} \end{bmatrix} = \begin{bmatrix} 2x_1 - 4 \\ 2x_2 - 5 \end{bmatrix}$$

Gradienten der Nebenbedingungen:

$$\nabla g_1(\mathbf{x}) = \begin{bmatrix} -2 \\ -1 \end{bmatrix} \quad \nabla g_2(\mathbf{x}) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \nabla g_3(\mathbf{x}) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

## Lösung des Problems:

**Step 1:** Ein zulässiger Schätzpunkt:

$$\mathbf{x}^0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad A^0 = \{2, 3\}, \quad k = 0$$

**Step 2:**

$$\min_{\mathbf{d}} \ \frac{1}{2}\mathbf{d}^T\mathbf{Q}\mathbf{d} + (\mathbf{Q}\mathbf{x}^k + \mathbf{c})^T\mathbf{d}$$

mit $\mathbf{a}_i^T\mathbf{d} = 0 \qquad i \in A^k$

Die Nebenbedingungen:

$$i = 2 \qquad \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = 0$$

$$i = 3 \qquad \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = 0$$

d. h. $d_1 = d_2 = 0$

**Step 3:** $\lambda$ prüfen durch

$$\mathbf{Q}\mathbf{x}^k + \mathbf{C} = \sum_i \lambda_i \mathbf{a}_i$$

d. h.

$$\begin{bmatrix} -4 \\ -5 \end{bmatrix} = \lambda_2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \lambda_3 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Ergibt sich

$$\lambda_2 = -4 \quad \lambda_3 = -5$$

dann

$$\lambda_j = min\{-4, -5\} = -5 \qquad j = 3$$

$$A^k = A^k - \{j\} = \{2\}$$

**Step 2:** Die Nebenbedingung:

$$i = 2 \qquad \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = 0 \qquad d_1 = 0$$

Das Optimierungsproblem:

$$min\ \frac{1}{2} \begin{bmatrix} 0 & d_2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ d_2 \end{bmatrix} + \begin{bmatrix} -4 & -5 \end{bmatrix} \begin{bmatrix} 0 \\ d_2 \end{bmatrix}$$

$$= min\ d_2^2 - 5d_2$$

Die Lösung $d_2 = 2,5$ und dann $\mathbf{d} = \begin{bmatrix} 0 \\ 2,5 \end{bmatrix}$

## Step 4:

$$\alpha_p = min\left\{ \frac{b_i - \mathbf{a}_i^T \mathbf{x}^k}{\mathbf{a}_i^T \mathbf{d}}, \quad i \notin A^k \right\}$$

$$i = 1 \qquad \mathbf{a}_1^T \mathbf{d} = \begin{bmatrix} -2 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 2,5 \end{bmatrix} = -2,5 < 0$$

$$\alpha = \frac{-2}{-2,5} = 0,8$$

$$i = 3 \qquad \mathbf{a}_3^T \mathbf{d} = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 2,5 \end{bmatrix} = 2,5 > 0$$

dann $\quad p = 1, \qquad A^{k+1} = A^k \cup \{1\} = \{1, 2\}$

**Step 5:** $k = 1$

$$\mathbf{x}^1 = \mathbf{x}^0 + \alpha \mathbf{d} = 0,8 \begin{bmatrix} 0 \\ 2,5 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$$

GOTO Step 2

## Die Lösung des Problems:

$$\mathbf{x}^* = \begin{bmatrix} 0,2 \\ 1,6 \end{bmatrix} \qquad \boldsymbol{\lambda}^* = \begin{bmatrix} 1,8 \\ 0 \\ 0 \end{bmatrix}$$

## Grafische Darstellung der Lösung:

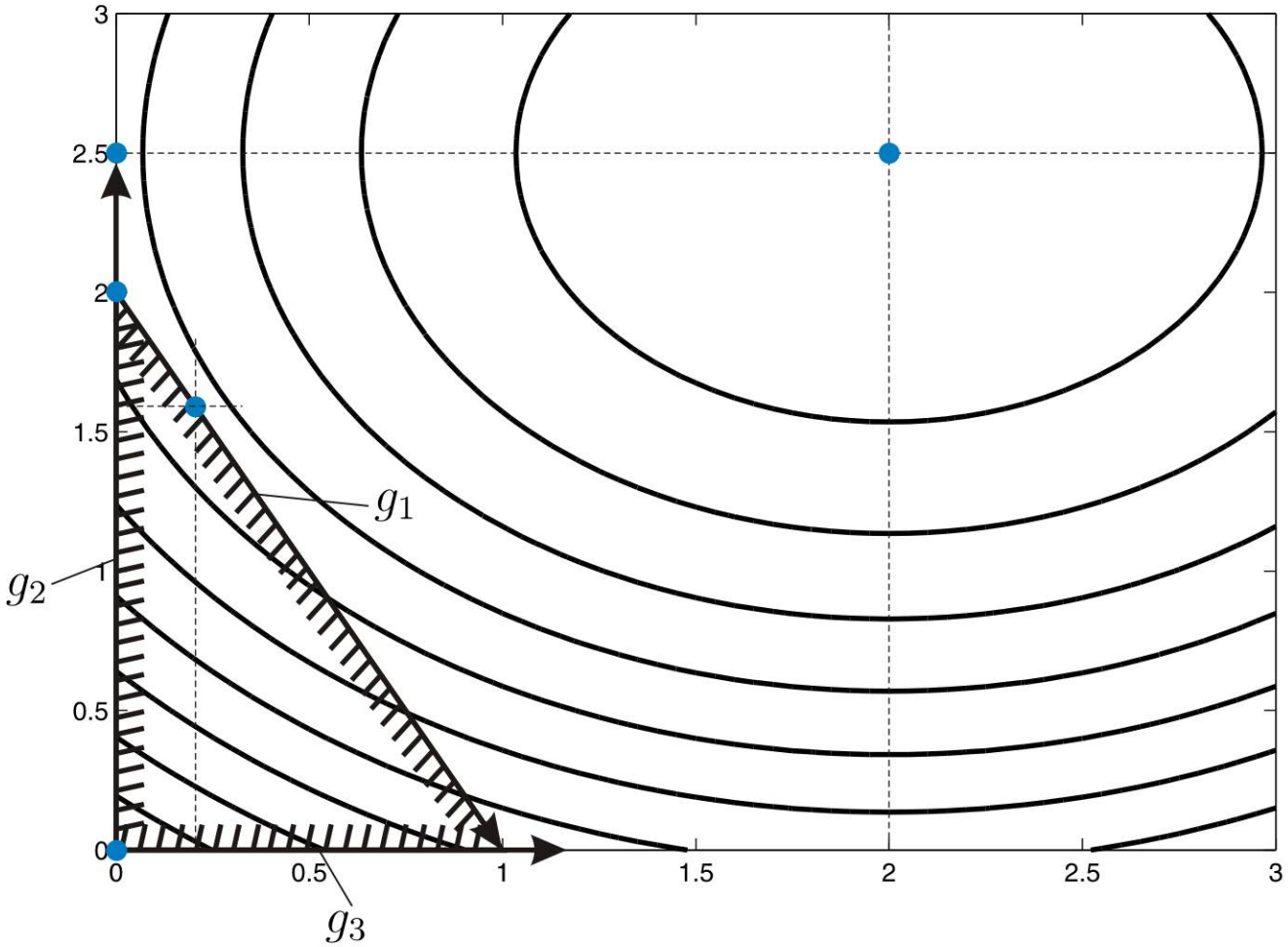
$$\min_{\mathbf{x}} \; f(\mathbf{x}) = x_1^2 + x_2^2 - 4x_1 - 5x_2 + 2$$

mit $g_1(\mathbf{x}) = -2x_1 - x_2 + 2 \geq 0$ (1)

$g_2(\mathbf{x}) = x_1 \geq 0$ (2)

$g_3(\mathbf{x}) = x_2 \geq 0$ (3)

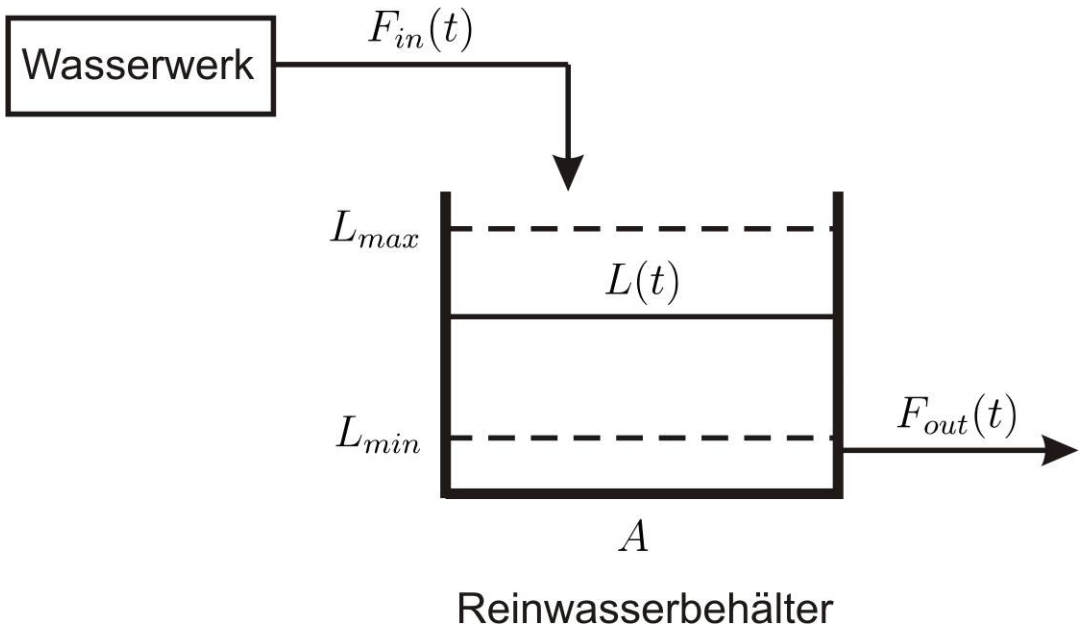
# Grafische Darstellung:



# Optimale Betriebsplanung eines Wasserwerks für 24 Stunden

## Diskrete Formulierung $(k = 1, \cdots, 24)$:

- Wasserbedarf $F_{out}(k)$ ist bekannt.
- Betriebskosten sind eine Funktion der Zeit $c(k)$, d. h. sie sind unterschiedlich am Tag und in der Nacht.
- Was ist die optimale Fahrweise für $F_{in}(k)$?



## Problemformulierung:

$$\min_{F_{in}} \sum_{k=1}^{24} c(k)F_{in}(k) + \beta \sum_{k=1}^{24} \left[F_{in}(k) - F_{in}(k-1)\right]^2$$

mit $$L(k+1) = L(k) + \frac{1}{A}\left[F_{in}(k) - F_{out}(k)\right]$$

$$L(24) = L(0) = L_0$$

$$L_{min} \leq L(k) \leq L_{max} \qquad k = 1, \cdots, 24$$

$$F_{min} \leq F_{in}(k) \leq F_{max}$$

## Parameter im Problem:

- Wasserbedarf ($m^3/h$): $F_{out}(k) = 500 - (16-k)^2$
- Kostenfunktion ($/$m^3$): $c(k) = \exp\lfloor -(k-12)^2/72 \rfloor$
- Tankfläche ($m^2$): $A = 200$
- Initialfüllstand (m): $L_0 = 5$
- Grenze des Füllstands (m): $L_{min} = 2, \; L_{max} = 8$
- Stromgrenze ($m^3/h$): $F_{in,min} = 200, \; F_{in,max} = 700$
- Gewichtsfaktor: $\beta = 0.01$

Fachgebiet Prozessoptimierung

## Variablen:

- Steuervariablen: $F_{in}(k) \quad k = 1, \cdots, 24$
- Zustandsvariablen: $L(k) \quad k = 1, \cdots, 24$

## Restriktionen:

### Gleichungsnebenbedingungen:

$$L(k+1) = L(k) + \frac{1}{A}\left[F_{in}(k) - F_{out}(k)\right] \quad k = 1, \cdots, 24$$

$$L(24) = L(0) = L_0$$

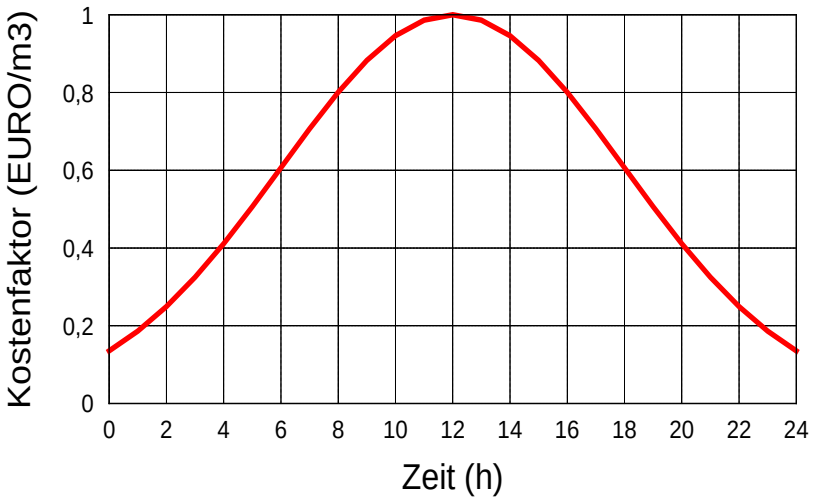
### Ungleichungsnebenbedingungen:

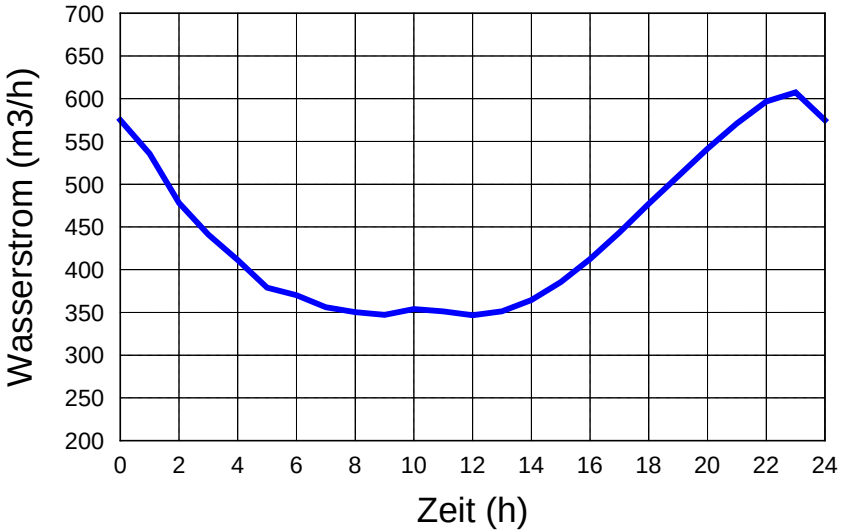
$$L_{min} \leq L(k) \leq L_{max} \qquad k = 1, \cdots, 24$$

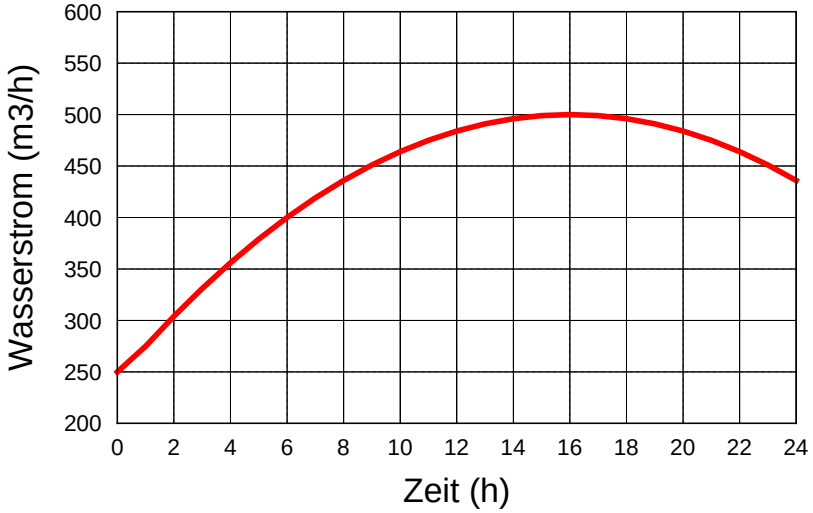
$$F_{min} \leq F_{in}(k) \leq F_{max} \qquad k = 1, \cdots, 24$$

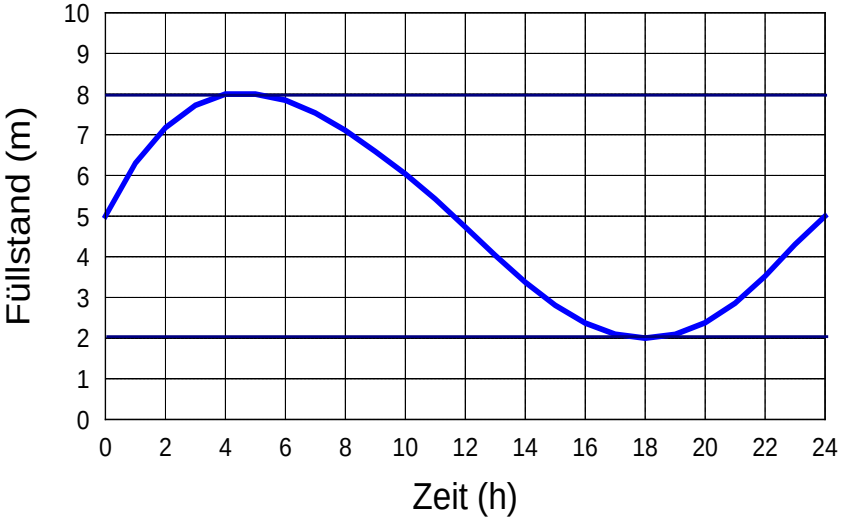
# Das Ergebnis Optimierung:






- Kosten beim optimalen Fahrplan: 5784.24 $
- Kosten beim konventionellen Fahrplan: 6558.07 $

Fachgebiet Prozessoptimierung

# Das SQP-Verfahren



Mit Gleichungsnebenbedingungen

$$\begin{aligned} min\ & f(\mathbf{x}) \\ & \mathbf{g}(\mathbf{x}) = \mathbf{0} \end{aligned}$$

Lagrange-Funktion:

$$L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) - \boldsymbol{\lambda}^T \mathbf{g}(\mathbf{x})$$

KKT-Bedingungen:

$$\begin{aligned} & \nabla L(\mathbf{x}, \boldsymbol{\lambda}) = \nabla f(\mathbf{x}) - [\nabla \mathbf{g}(\mathbf{x})]^T \boldsymbol{\lambda} = \mathbf{0} \\ & \mathbf{g}(\mathbf{x}) = \mathbf{0} \end{aligned}$$

Taylor-Entwicklung erster Ordnung von: $\mathbf{x}^k, \boldsymbol{\lambda}^k$

$$\nabla L(\mathbf{x}^k, \boldsymbol{\lambda}^k) + \nabla^2 L(\mathbf{x}^k, \boldsymbol{\lambda}^k) \Delta \mathbf{x}^k - \left[\nabla \mathbf{g}(\mathbf{x}^k)\right]^T \Delta \boldsymbol{\lambda}^k = \mathbf{0}$$

d. h.

$$\nabla f(\mathbf{x}^k) - \left[\nabla \mathbf{g}(\mathbf{x}^k)\right]^T \boldsymbol{\lambda}^k + \nabla_{\mathbf{x}}^2 L(\mathbf{x}^k, \boldsymbol{\lambda}^k)\Delta\mathbf{x}^k - \left[\nabla \mathbf{g}(\mathbf{x}^k)\right]^T \Delta\boldsymbol{\lambda}^k = \mathbf{0}$$

und

$$\mathbf{g}(\mathbf{x}^k) + \left[\nabla \mathbf{g}(\mathbf{x}^k)\right]^T \Delta\mathbf{x}^k = \mathbf{0}$$

d. h.

$$\begin{bmatrix} \nabla^2 L(\mathbf{x}^k, \boldsymbol{\lambda}^k) & \left[\nabla \mathbf{g}(\mathbf{x}^k)\right]^T \\ \nabla \mathbf{g}(\mathbf{x}^k) & 0 \end{bmatrix} \begin{bmatrix} \Delta\mathbf{x}^k \\ -\Delta\boldsymbol{\lambda}^k \end{bmatrix} = \begin{bmatrix} -\nabla f(\mathbf{x}^k) + \left[\nabla \mathbf{g}(\mathbf{x}^k)\right]^T \boldsymbol{\lambda}^k \\ -\mathbf{g}(\mathbf{x}^k) \end{bmatrix}$$

Die Lösung des Gleichungssystems ist die Lösung des folgenden QPE-Problems mit $\mathbf{d} = \Delta\mathbf{x}^k$

$$\min_{\mathbf{d}} \ \frac{1}{2}\mathbf{d}^{\mathbf{T}}\left[\nabla^2 L(\mathbf{x}^k, \boldsymbol{\lambda}^k)\right]\mathbf{d} + \left[\nabla f(\mathbf{x}^k)\right]^T \mathbf{d}$$

mit $\nabla \mathbf{g}(\mathbf{x}^k)\mathbf{d} + \mathbf{g}(\mathbf{x}^k) = \mathbf{0}$

Die Lagrange-Funktion des Problems:

$$\tilde{L}(\mathbf{d}, \boldsymbol{\mu}) = \frac{1}{2}\mathbf{d}^T L_{xx}^2(\mathbf{x}^k, \boldsymbol{\lambda}^k)\mathbf{d} + \left[\nabla f(\mathbf{x}^k)\right]^T \mathbf{d} - \boldsymbol{\mu}^T\left\{\mathbf{g}(\mathbf{x}^k) + \nabla \mathbf{g}(\mathbf{x}^k)\mathbf{d}\right\}$$

KKT-Bedingungen:

$$\nabla_{\mathbf{d}}\tilde{L}(\mathbf{d}, \boldsymbol{\mu}) = L_{xx}^2(\mathbf{x}^k, \boldsymbol{\lambda}^k)\mathbf{d} + \nabla f(\mathbf{x}^k) - \left[\nabla \mathbf{g}(\mathbf{x}^k)\right]^T \boldsymbol{\mu} = \mathbf{0}$$

$$\mathbf{g}(\mathbf{x}^k) + \left[\nabla \mathbf{g}(\mathbf{x}^k)\right]^T \mathbf{d} = \mathbf{0}$$

nun setzt man

$$\mathbf{d} = \Delta \mathbf{x}, \quad \boldsymbol{\mu} = \Delta\boldsymbol{\lambda} + \boldsymbol{\lambda}^k = \boldsymbol{\lambda}^{k+1}$$

dann

$$\nabla\tilde{L}(\mathbf{d}, \boldsymbol{\mu}) = L_{xx}^2(\mathbf{x}^k, \boldsymbol{\lambda}^k)\Delta\mathbf{x} + \nabla \mathbf{f}(\mathbf{x}^k) - \left[\nabla \mathbf{g}(\mathbf{x}^k)\right]^T(\Delta\boldsymbol{\lambda} + \boldsymbol{\lambda}^k)$$

$$= L_{xx}^2(\mathbf{x}^k, \boldsymbol{\lambda}^k)\Delta\mathbf{x} + \nabla f(\mathbf{x}^k) - \left[\nabla \mathbf{g}(\mathbf{x}^k)\right]^T\boldsymbol{\lambda}^k - \left[\nabla \mathbf{g}(\mathbf{x}^k)\right]^T\Delta\boldsymbol{\lambda}$$

$$= L_{xx}^2(\mathbf{x}^k, \boldsymbol{\lambda}^k)\Delta\mathbf{x} + \nabla L(\mathbf{x}^k, \boldsymbol{\lambda}^k) - \left[\nabla \mathbf{g}(\mathbf{x}^k)\right]^T\Delta\boldsymbol{\lambda} = \mathbf{0}$$

und $$\mathbf{g}(\mathbf{x}^k) + \nabla\mathbf{g}(\mathbf{x}^k)\Delta\mathbf{x} = \mathbf{0}$$

d. h. beim Problem mit Gleíchungsnebenbedingungen kann man eine Änderung $\mathbf{d}$ durch Lösung des folgenden Problems erhalten:

$$\begin{aligned} \min_{\mathbf{d}} \quad & \frac{1}{2}\mathbf{d}^{\mathbf{T}} L^2_{xx}(\mathbf{x}^k, \boldsymbol{\lambda}^k)\mathbf{d} + \left[\nabla f(\mathbf{x}^k)\right]^T \mathbf{d} \\ \text{mit} \quad & \mathbf{g}(\mathbf{x}^k) + \nabla\mathbf{g}(\mathbf{x}^k)\mathbf{d} = \mathbf{0} \\ & \mathbf{x}^{k+1} = \mathbf{x}^k + \mathbf{d} \end{aligned}$$

**Beispiel:**

$$\begin{aligned} \min_{\mathbf{x}} \quad & f(\mathbf{x}) = (x_1 - 5)^2 + x_2^2 \\ \text{mit} \quad & g(\mathbf{x}) = x_2 - x_1^2 = 0 \end{aligned}$$

Die Lagrange-Funktion:

$$L(\mathbf{x}, \lambda) = (x_1 - 5)^2 + x_2^2 - \lambda(x_2 - x_1^2)$$

Die Hesse-Matrix von der Lagrange-Funktion und der Gradientenvektor von der Zielfunktion:

$$L_{xx}^2(\mathbf{x}^k, \boldsymbol{\lambda}^k) = \begin{bmatrix} 2+2\lambda^k & 0 \\ 0 & 2 \end{bmatrix} \qquad \nabla f(\mathbf{x}^k) = \begin{bmatrix} 2(x_1^k - 5) & 2x_2^k \end{bmatrix}^T$$

Das QPE-Problem:

$$\min_{d_1,d_2} \frac{1}{2} \begin{bmatrix} d_1 & d_2 \end{bmatrix} \begin{bmatrix} 2+2\lambda^k & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} + \begin{bmatrix} 2(x_1^k - 5) & 2x_2^k \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$$

mit $\begin{bmatrix} -2x_1^k & 1 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} + x_2^k - (x_1^k)^2 = 0$

$$\left[\nabla \mathbf{g}(\mathbf{x}^k)\right]^T \mathbf{d} + \mathbf{g}(\mathbf{x}^k) = \mathbf{0}$$

Die Lagrange-Funktion dieses Problems:

$$\begin{aligned} \tilde{L}(d_1, d_2, \mu) = & (1+\lambda^k)d_1^2 + d_2^2 + 2(x_1^k - 5)d_1 + 2x_2^k d_2 \\ & - \mu[-2x_1^k d_1 + d_2 + x_2^k - (x_1^k)^2] \end{aligned}$$

Nach KKT-Bedingungen:

$$\frac{\partial \tilde{L}}{\partial d_1} = 2(1+\lambda^k)d_1 + 2(x_1^k - 5) + 2\mu\, x_1^k = 0$$

$$\frac{\partial \tilde{L}}{\partial d_2} = 2d_2 + 2x_2^k - \mu = 0$$

$$-2x_1^k d_1 + d_2 + x_2^k - (x_1^k)^2 = 0$$

Fängt man mit $x_1^0 = x_2^0 = 5, \quad \lambda^0 = 0$ an:

| $k$ | $x_1^k$ | $x_2^k$ | $\lambda^k$ | $d_1^k$ | $d_2^k$ | $f^k$ | $g^k$ |
|---|---|---|---|---|---|---|---|
| 1 | 5.000000 | 5.000000 | 0.000000 | −2.475248 | −4.752476 | 25.000000 | −20.000000 |
| 2 | 2.524752 | 0.247524 | 0.495049 | −1.100755 | 0.568582 | 6.188119 | −6.126850 |
| 3 | 1.423997 | 0.816106 | 1.632212 | −0.204692 | 0.628699 | 13.453825 | −1.211662 |
| 4 | 1.219305 | 1.444805 | 2.889610 | 0.015779 | 0.080377 | 16.381119 | −0.041899 |
| 5 | 1.235083 | 1.525182 | 3.050364 | −0.000310 | −0.000518 | 16.500778 | −0.000249 |
| 6 | 1.234773 | 1.524664 | 3.049328 | 0.000000 | 0.000000 | 16.501535 | 0.000000 |

## Merit Function

Die Suchrichtung in jeder Iteration $\mathbf{p}^k$ zur Reduktion der Zielfunktion, aber beim neuen Punkt werden die Nebenbedingungen verletzt.

## Wie groß soll die Schrittlänge sein?

Ein Kompromiss zwischen dem Zielfunktionsabstieg und der Zulässigkeit wird durch Line-Search gefunden.

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \alpha \mathbf{d}^k$$

Man definiert sog. Merit Function (Straffunktion):

$$\min_{\mathbf{x}} \ \varphi(\mathbf{x}^{k+1}) = \varphi(\mathbf{x}^k + \alpha \mathbf{d}^k)$$
$$= f(\mathbf{x}^k + \alpha \mathbf{d}^k) + \frac{1}{\mu} \sum_{j=1}^{m} \left| g_j(\mathbf{x}^k + \alpha \mathbf{d}^k) \right|$$

oder

$$\min_{\mathbf{x}} \ \varphi(\mathbf{x}^{k+1}) = f(\mathbf{x}^k + \alpha \mathbf{d}^k) - \sum_{j=1}^{m} \lambda_j g_j(\mathbf{x}^k + \alpha \mathbf{d}^k)$$
$$+ \frac{1}{\mu} \sum_{j=1}^{m} \left[ g_j(\mathbf{x}^k + \alpha \mathbf{d}^k) \right]^2$$

**Problem:** der Gewichtungsfaktor ist schwer zu definieren.

# Algorithmus Line-Search-SQP (Sequentielle Quadratische Programmierung)

## Nur mit Gleichungsnebenbedingungen

Schritt 1: $k = 0$, Vorgabe Schätzwerte $\mathbf{x}^0, \boldsymbol{\lambda}^0$ ($\mathbf{x}^0$ nicht unbedingt zulässig).

Schritt 2: In Iteration $k$, Berechnung des Gradientenvektors $\nabla f(\mathbf{x}^k)$ und Approximation der Hesse-Matrix $\mathbf{H}_L(\mathbf{x}^k, \boldsymbol{\lambda}^k)$ mit der BFGS-Formel.

Schritt 3: Lösung des QPE-Problems

$$\min_{\mathbf{d}} \ \frac{1}{2}\mathbf{d}^{\mathbf{T}}\mathbf{H}_L(\mathbf{x}^k, \boldsymbol{\lambda}^k)\mathbf{d} + \left[\nabla f(\mathbf{x}^k)\right]^T \mathbf{d}$$

mit $\mathbf{g}(\mathbf{x}^k) + \left[\nabla \mathbf{g}(\mathbf{x}^k)\right]^T \mathbf{d} = \mathbf{0}$

Durch analytische Lösung erhält man $\mathbf{d}^*, \boldsymbol{\mu}^*$.

Schritt 4: Wenn die Optimalitätsbedingungen erfüllt sind, STOP. Ansonsten geht zu Schritt 5.

Schritt 5: Line Search zur Ermittlung $\alpha$.

Schritt 6: $\mathbf{x}^{k+1} = \mathbf{x}^k + \alpha \mathbf{d}^*, \boldsymbol{\lambda}^{k+1} = \boldsymbol{\mu}^*$

Schritt 7: $k \leftarrow k + 1$, Fortschritt mit Schritt 2.

# Sequentielle quadratische Programmierung (das SQP-Verfahren)

Problemdarstellung:

$$\min_{\mathbf{x}} \; f(\mathbf{x})$$

$$\text{mit} \quad g_i(\mathbf{x}) = 0 \qquad i = 1, \cdots, m$$

$$g_i(\mathbf{x}) \geq 0 \qquad i = m+1, \cdots, m+r$$

$$x_{min\,j} \leq x_j \leq x_{max\,j} \qquad j = 1, \cdots, n$$

$f(\mathbf{x}), g_i(\mathbf{x})$ sind nichtlineare Funktionen, zusätzlich

$$g_{m+r+j} = x_{max\,j} - x_j \geq 0 \qquad j = 1, \cdots, n$$

$$g_{m+r+n+j} = x_j - x_{min\,j} \geq 0 \qquad j = 1, \cdots, n$$

Lagrange-Funktion:

$$L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) - \sum_{i=1}^{m+r+2n} \lambda_i g_i(\mathbf{x})$$

## Lösungsverfahren:

Taylor-Entwicklung von $\mathbf{x}^k$

$$L(\mathbf{x}^k + \mathbf{d}) = f(\mathbf{x}^k) + \left[\nabla f(\mathbf{x}^k)\right]^T \mathbf{d} + \frac{1}{2}\mathbf{d}^T \mathbf{H}_L(\mathbf{x}^k, \boldsymbol{\lambda}^k)\mathbf{d}$$

$$g_i(\mathbf{x}^k + \mathbf{d}) = g_i(\mathbf{x}^k) + \left[\nabla f(\mathbf{x}^k)\right]^T \mathbf{d}\left[\nabla g_i(\mathbf{x}^k)\right]^T \mathbf{d}$$

$\mathbf{H}(\mathbf{x}^k, \boldsymbol{\lambda}^k)$ ist eine positive definite Approximation der Hesse-Matrix von $L(\mathbf{x}, \boldsymbol{\lambda})$ .

Nun sieht das Optimierungsproblem wie folgt aus:

$$\min_{\mathbf{d}} \frac{1}{2}\mathbf{d}^T \mathbf{H}(\mathbf{x}^k, \boldsymbol{\lambda}^k)\mathbf{d} + \left[\nabla f(\mathbf{x}^k)\right]^T \mathbf{d}$$

mit $g_i(\mathbf{x}^k) + \nabla g_i(\mathbf{x}^k)\mathbf{d} = 0 \qquad i = 1, \cdots, m$

$g_i(\mathbf{x}^k) + \nabla g_i(\mathbf{x}^k)\mathbf{d} \geq 0 \qquad i = m+1, \cdots, m+r+2n$

Dieses Problem kann mit dem „active set“ Verfahren gelöst werden. In Iteration $k$ werden mit $\mathbf{x}^k, \boldsymbol{\lambda}^k, \mathbf{H}^k$ für Iteration $k+1$ $\mathbf{x}^{k+1}, \boldsymbol{\lambda}^{k+1}$ berechnet. Die Hesse-Matrix der Langrange-Funktion wird mit dem BFGS-Ansatz approximiert.

## Das SQP-Verfahren:

- $f(\mathbf{x})$ mit einer quadratischen Funktion annähern
- $g(\mathbf{x})$ mit linearen Funktionen annähern
- Das Problem mit dem „active set“ Verfahren lösen.
- Aufgrund der Annäherung treten Konvergenzprobleme auf.

# Algorithmus Active-Set-SQP (Sequentielle Quadratische Programmierung)

## Mit Gleichungs- und Ungleichungsnebenbedingungen

Schritt 1: $k = 0$, Vorgabe Schätzwerte $\mathbf{x}^0, \boldsymbol{\lambda}^0$,
$\mathbf{x}^0$ muss zulässig sein. Identifikation Active-Set $A^0$.

Schritt 2: In Iteration $k$, Berechnung des Gradientenvektors $\nabla f(\mathbf{x}^k)$ und Approximation der Hesse-Matrix $\mathbf{H}_L(\mathbf{x}^k, \boldsymbol{\lambda}^k)$ mit der BFGS-Formel.

Schritt 3: Iterative Lösung des QP-Problems

$$\min_{\mathbf{d}} \ \frac{1}{2}\mathbf{d}^{\mathbf{T}}\mathbf{H}_L(\mathbf{x}^k, \boldsymbol{\lambda}^k)\mathbf{d} + \left[\nabla f(\mathbf{x}^k)\right]^T \mathbf{d}$$

mit
$$\mathbf{g}_i(\mathbf{x}^k) + \left[\nabla \mathbf{g}_i(\mathbf{x}^k)\right]^T \mathbf{d} = \mathbf{0} \qquad i \in A^k$$
$$\mathbf{g}_i(\mathbf{x}^k) + \left[\nabla \mathbf{g}_i(\mathbf{x}^k)\right]^T \mathbf{d} \geq \mathbf{0} \qquad i \notin A^k$$

mit dem Active-Set-Verfahren. Man erhält $\mathbf{d}^*, \boldsymbol{\mu}^*$.

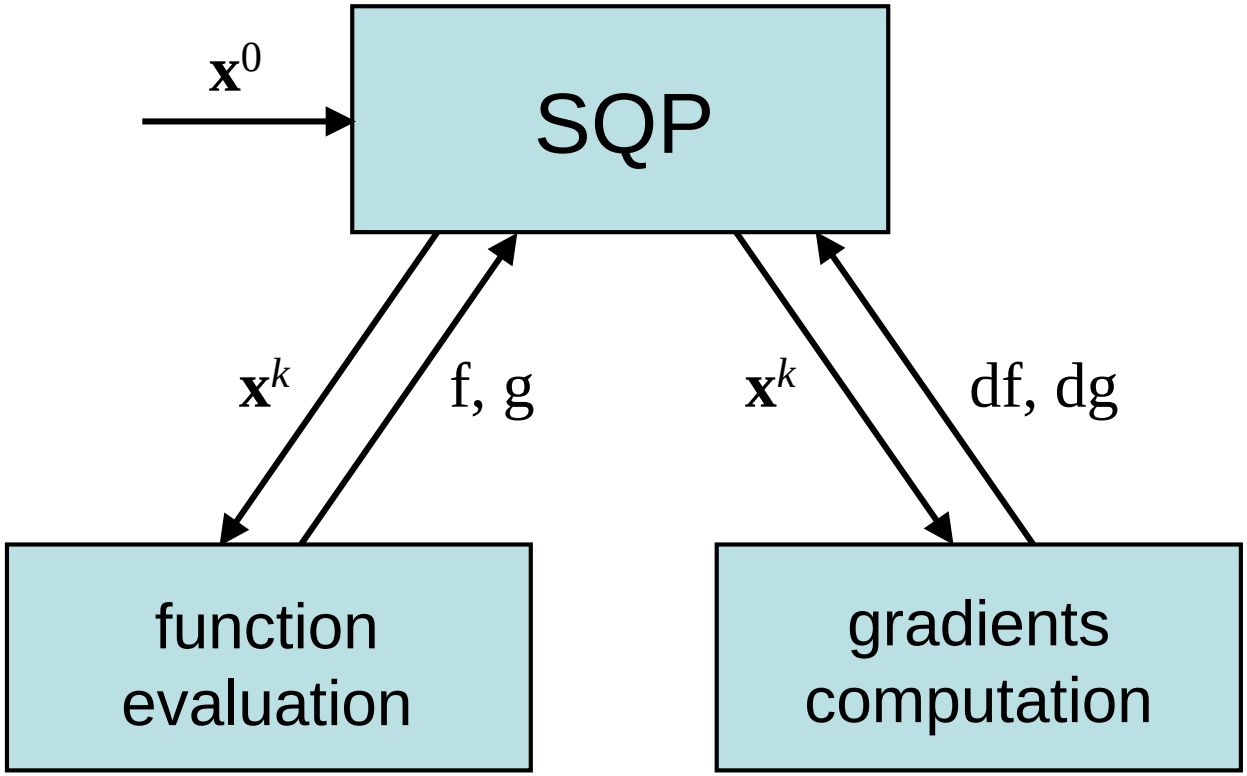
Schritt 4: Überprüfung der nicht aktiven Ungleichungen.
Ermittlung der Schrittlänge $\alpha$ .

Schritt 5: $\mathbf{x}^{k+1} = \mathbf{x}^k + \alpha \mathbf{d}^*, \boldsymbol{\lambda}^{k+1} = \boldsymbol{\mu}^*$ , Update Active-Set.

Schritt 6: Wenn die Optimalitätsbedingungen erfüllt sind, STOP.
Ansonsten zu Schritt 7.

Schritt 7: $k \leftarrow k + 1$, Fortschritt mit Schritt 2.

# Computation strategy for solving NLP problems



## Vorhandene NLP Software

- NLPQL, Schittkowski (1986), *Annals of Operations Research*, Vol. 5, S. 485-500.
- MINOS, Murtagh und Saunders (1978), *Mathematical Programming*, Vol. 14, S. 41-72.
- GINO, Lieman et al. (1986), *Modelling and Optimization with GINO*, Redwood City, Scientific Press.
- GAMS, Brooke et al. (1988), *GAMS-A User Guide*, Redwood City, Scientific Press.
- IMSL, (1987) Math/Library, IMSL User`s Manual.
- **SNOPT, Gill, Stanford University**
- **IPOPT, Biegler, Carnegie Mellon University**

## Anwendungsbeispiele:

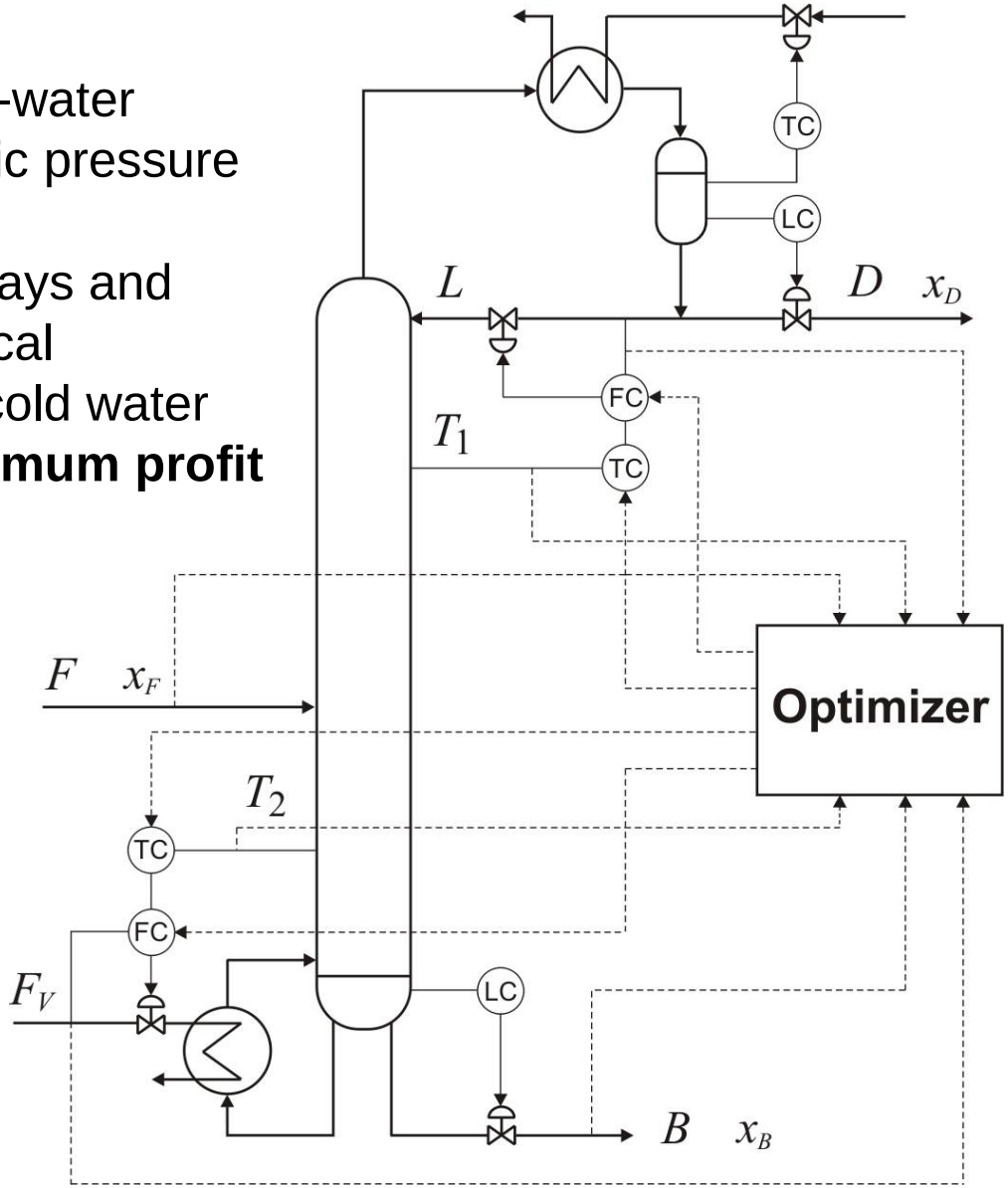
Morari und Grossmann (1991), Chemical Engineering Optimization Problems with GAMS, *CACHE Design Case Studies*, Vol. 6.

Biegler et al. (1997), Systematic Methods of Chemical Process Design, Prentice Hall.

# On-line Optimization of a distillation column

Separation of a methanol-water mixture under atmospheric pressure

Column: 20 bubble cap trays and 100 mm diameter. Electrical reboiler, condenser with cold water
Aim of optimization: **maximum profit**




Fachgebiet Prozessoptimierung

## Objective function:

$$\max_{L,F_V} \quad c_1 D x_D + c_2 B x_B - c_3 F x_F - c_4 F_V - c_5 F_W$$

## Equality constrainsts for each tray:

- Component balances
- Vapor-liquid equilibrium relations
- Energy balance
- Tray hydraulics

## Inequality constrainsts:

- State variable constrainst
- Control variable constrainst

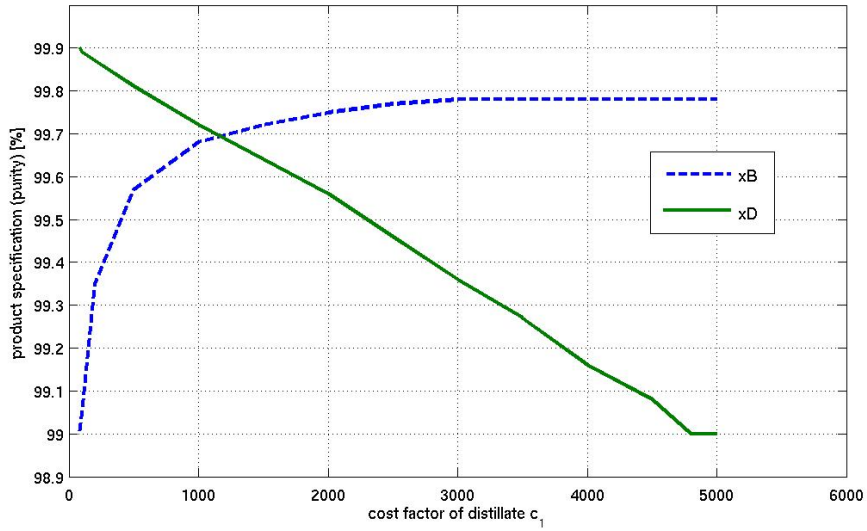
## Influential factors on optimality:

- Price factors: $c_1, c_2, c_3, c_4, c_5$
- Feed conditions: $F, x_F$
- Product specifications: $D, B, x_D, x_B$

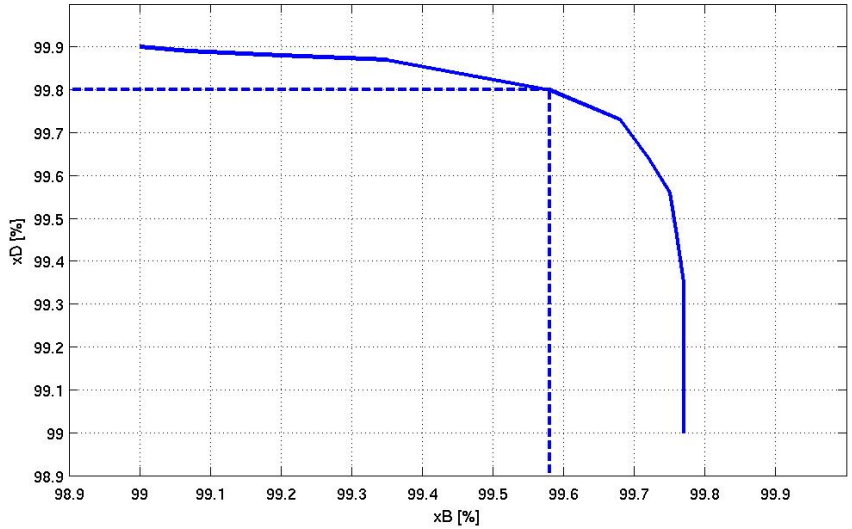
## Degree of freedom: 2

- $x_D$, $x_B$
- or $L$, $Q$

# Optimal operating point changing based on $C_1$

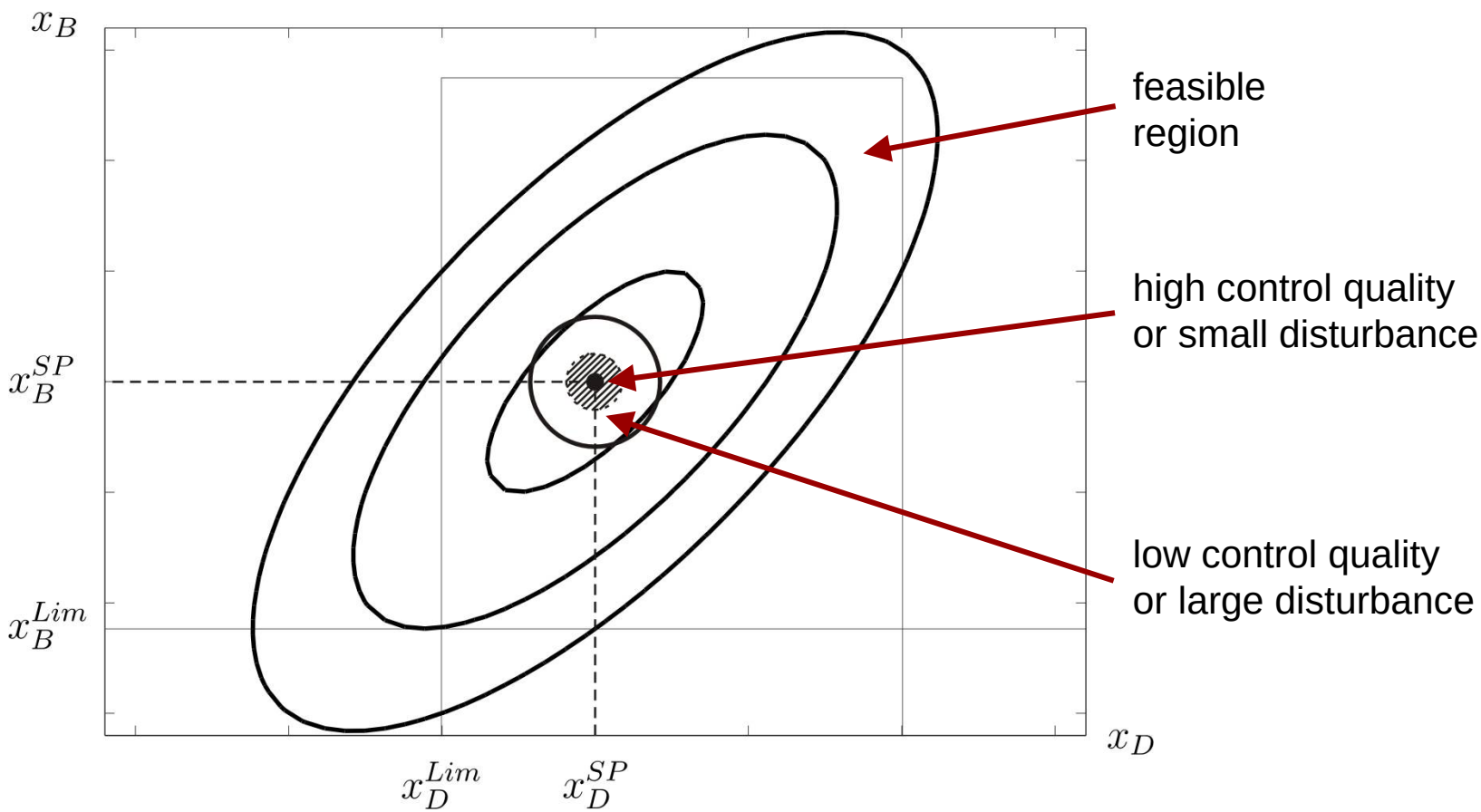



Influence of $C_1$



# Relation between optimization and control

## Minimum point is inside the feasible region



# Minimum point is outside the feasible region