Statische Prozessoptimierung/ Prozessoptimierung 1

Kapitel 4: Konvexitätsanalyse nichtlinearer Optimierungsprobleme

Prof. Dr.-Ing. habil. Pu Li

Fachgebiet Prozessoptimierung

Optimierung nichtlinearer Systeme

Problemdarstellung

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

$$\text{mit} \quad \mathbf{g}(\mathbf{x}) = 0$$

$$\mathbf{h}(\mathbf{x}) \leq 0$$

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

$f(\mathbf{x})$, $\mathbf{g}(\mathbf{x})$, $\mathbf{h}(\mathbf{x})$ sind nichtlineare Funktionen.

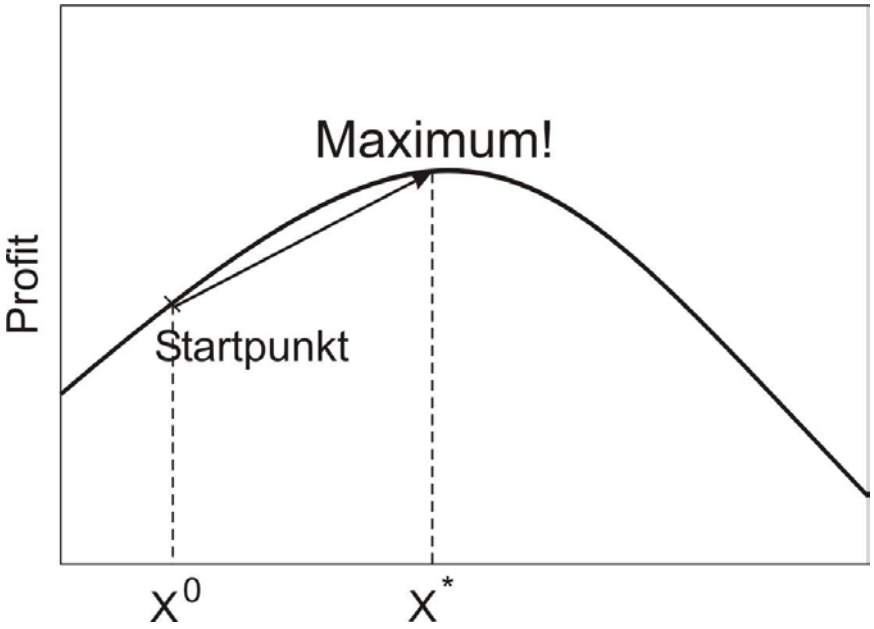
Charakter des Problems:

- Die Zielfunktion ist oft nicht monoton.
- Die Nebenbedingungen bilden einen komplexen zulässigen Bereich.
- Häufig gibt es lokale Lösungen.
- Das globale Optimum soll gefunden werden.

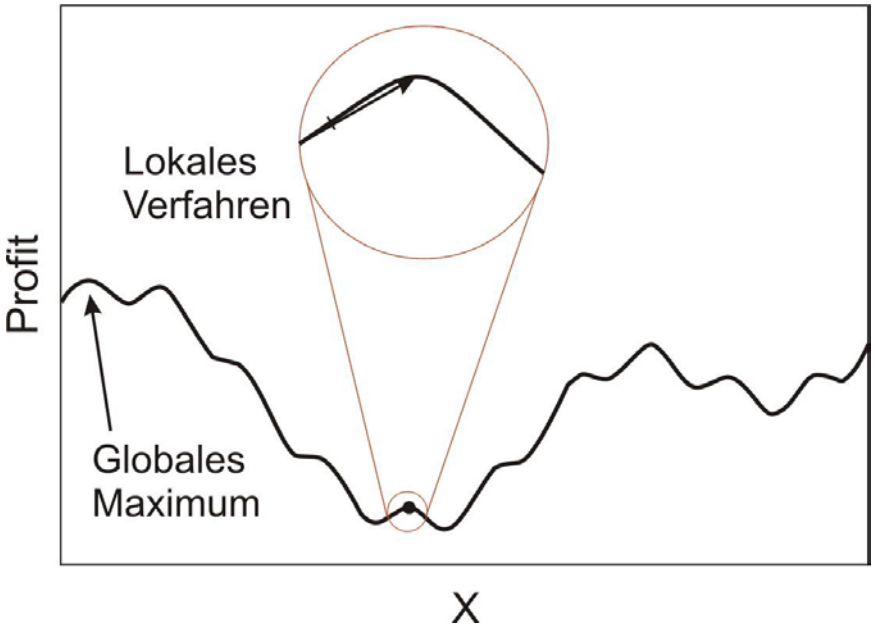
Analyse der Konvexität des Problems:

- Die Zielfunktion
- Der zulässige Bereich

Ein lokales Maximum



Finden des globalen Maximums



Eine konvexe Funktion $f(x)$

Zwei beliebige Punkte: $x_1,\ x_2 \subset C$ und $0 \leq \alpha \leq 1$

Es gibt:

$$x(\alpha) = \alpha\, x_1 + (1-\alpha)x_2$$

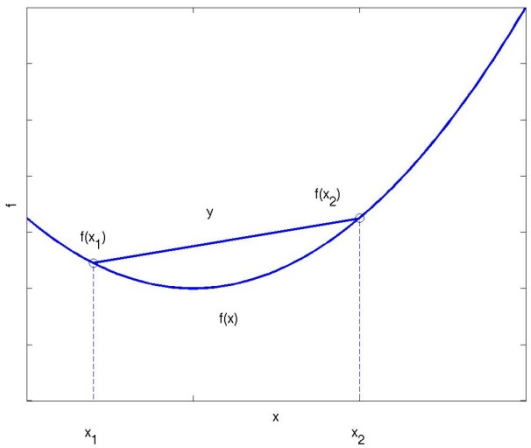
$$f(x) = f(\alpha\, x_1 + (1-\alpha)x_2)$$

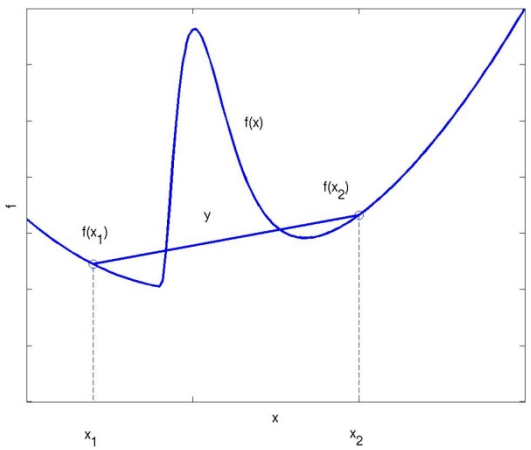
$$y = \alpha f(x_1) + (1-\alpha) f(x_2)$$

und

$$y \geq f(x)$$

dann ist die Funktion $f(x)$ konvex. Sonst ist sie nichtkonvex.

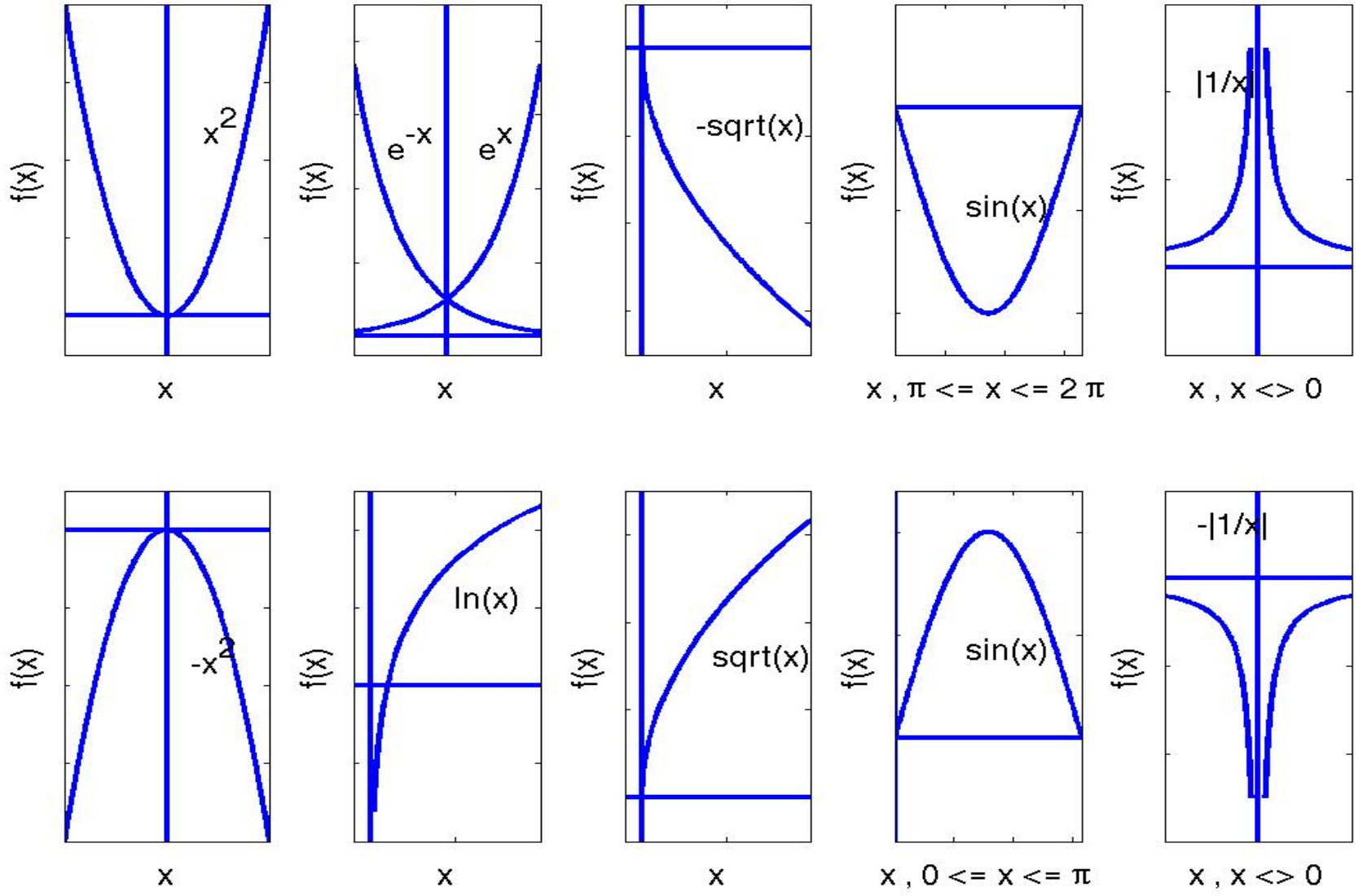




- Wenn $y > f(x)$ und $0 < \alpha < 1$, dann ist $f(x)$ „streng konvex“.
  Eine lineare Funktion ist konvex, aber nicht streng konvex.

- Wenn $f(x)$ konvex ist, ist auch $\lambda f(x)$ $(\lambda > 0)$ konvex.
- Wenn $f_1(x)$ und $f_2(x)$ konvex sind, ist auch $f_1(x) + f_2(x)$ konvex.

Gewöhnliche konvexe und konkave Funktionen



Taylor-Entwicklung einer Funktion:

Von einem Punkt: $\mathbf{x}^0$

$$f(\mathbf{x}) = f(\mathbf{x}^0) + [\nabla f(\mathbf{x}^0)]^T \Delta\mathbf{x} + \frac{1}{2}(\Delta\mathbf{x})^T \mathbf{H}(\mathbf{x}^0)\Delta\mathbf{x} + O(\Delta\mathbf{x}^3)$$

Eine Funktion mit zwei Variablen:

$$f(x_1, x_2) \approx f(x_1^0, x_2^0) + \begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} \end{bmatrix}_0^T \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \end{bmatrix}$$
$$+ \frac{1}{2} \begin{bmatrix} \Delta x_1 & \Delta x_2 \end{bmatrix} \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix}_0 \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \end{bmatrix}$$

Eine Zielfunktion mit mehreren Variablen

$$f(\mathbf{x}) = a_1 x_1^2 + a_2 x_2^2 + a_3 x_1 x_2 + a_4 x_1 + a_5 x_2 + a_6$$

die ersten Ableitungen (Gradienten):

$$\nabla f(\mathbf{x}) = \begin{bmatrix} \dfrac{\partial f}{\partial x_1} \\ \dfrac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 2a_1 x_1 + a_3 x_2 + a_4 \\ 2a_2 x_2 + a_3 x_1 + a_5 \end{bmatrix}$$

die zweiten Ableitungen (Hesse-Matrix):

$$\mathbf{H} = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} \\ \dfrac{\partial^2 f}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} 2a & a_3 \\ a_3 & 2a_2 \end{bmatrix}$$

die Taylor-Entwicklung:

$$f(\mathbf{x}) \stackrel{!}{=} f(\mathbf{x}^0) + \nabla f(\mathbf{x}^0)^T (\mathbf{x} - \mathbf{x}^0) + \frac{1}{2} (\mathbf{x} - \mathbf{x}^0)^T \mathbf{H}(\mathbf{x}^0)(\mathbf{x} - \mathbf{x}^0)$$

$\mathbf{H}$ positiv semidefinit $\Rightarrow$ $f(\mathbf{x})$ am Punkt $\mathbf{x}^0$ konvex

$\mathbf{H}$ negativ semidefinit $\Rightarrow$ $f(\mathbf{x})$ am Punkt $\mathbf{x}^0$ konkav

$\mathbf{H}$ indefinit $\Rightarrow$ $f(\mathbf{x})$ am Punkt $\mathbf{x}^0$ weder konvex noch konkav

Erste und zweite Ableitungen:

Ein Punkt $\hat{x}$ ist ein **stationärer Punkt** der Funktion $f(x)$, wenn der Gradientenvektor an diesem Punkt ein Nullvektor ist.

$$\nabla f(\hat{\mathbf{x}}) = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \cdots & \frac{\partial f}{\partial x_n} \end{bmatrix}^T = \mathbf{0}$$

Eine Funktion $f(x)$ ist am Punkt $\hat{x}$ **konvex**, wenn die Hesse-Matrix $\mathbf{H}$ der Funktion an diesem Punkt **positiv semidefinit** ist, nämlich

$$\mathbf{p}^T \mathbf{H}(\hat{\mathbf{x}}) \mathbf{p} \geq 0$$

wobei $\mathbf{p}$ ein beliebiger Vektor mit n Elementen ist, und

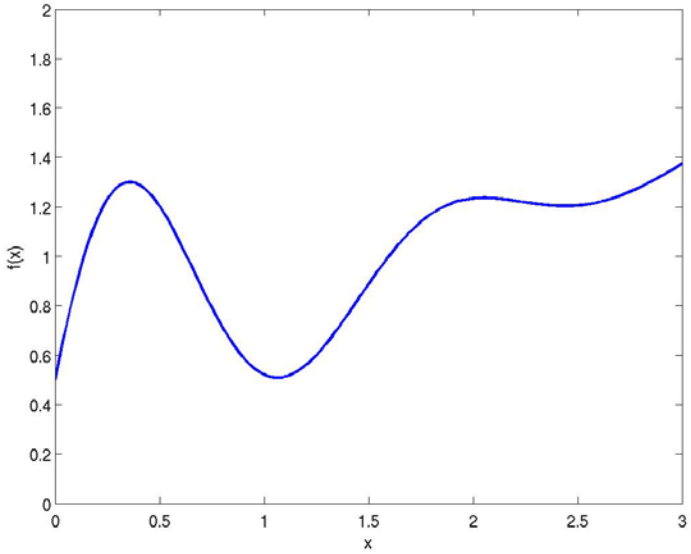
$$\mathbf{H}(\mathbf{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \cdots & \ddots & \cdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

Lokal oder global?

- Wenn $\hat{x}$ ein stationärer Punkt von $f(x)$ ist, dann hat die Funktion an diesem Punkt ein lokales Minimum/Maximum.
- Wenn die Hesse-Matrix der Funktion überall positiv definit ist, ist das lokale Minimum auch das globale Minimum.
- Wenn die Hesse-Matrix der Funktion negativ definit ist, ist die Funktion konkav. Dann gibt es mehrere lokale Minima.
- Es ist zu beachten, dass die Hesse-Matrix eine symmetrische Matrix aber nicht unbedingt überall positiv semidefinit ist.

Beispiel: eine nichtlineare Funktion:

$$f(x) = 0.5 + 0.3\,x + e^{-x} sin(4x)$$

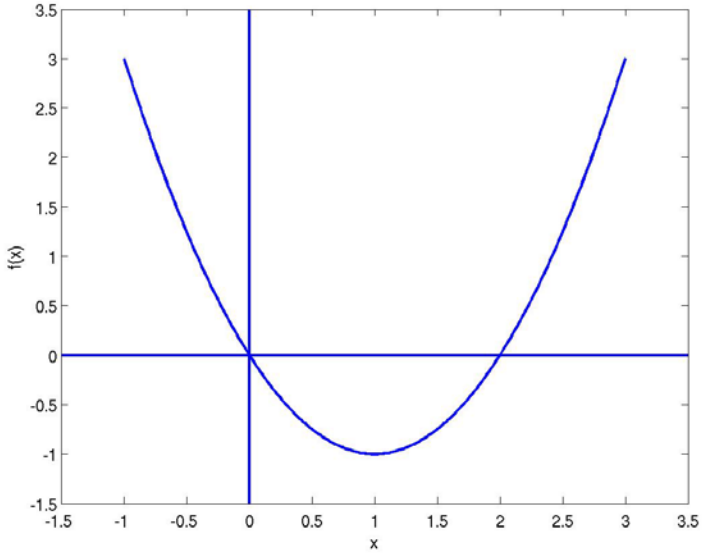


Beispiel: Konvex oder Konkav?

(1) $f(x) = x^2 - 2\,x$

$f'(x) = 2\,x - 2$

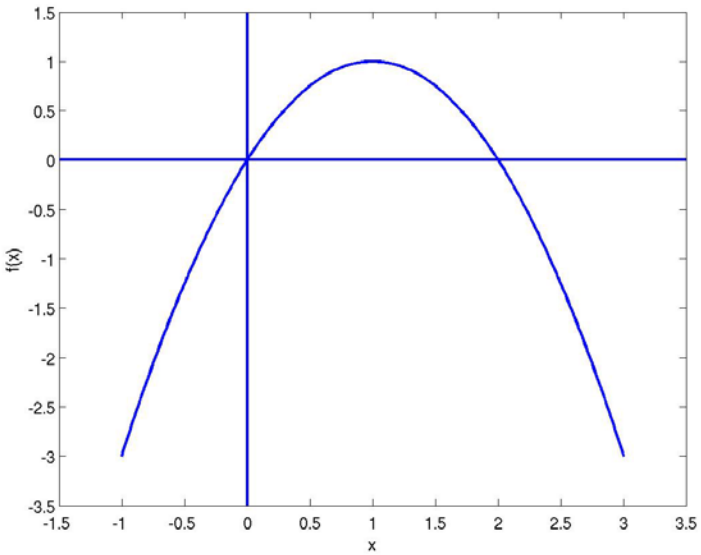
$f''(x) = 2 > 0,$ konvex



(2) $f(x) = -x^2 - 2\,x$

$f'(x) = -2\,x + 2$

$f''(x) = -2 < 0,$ konkav



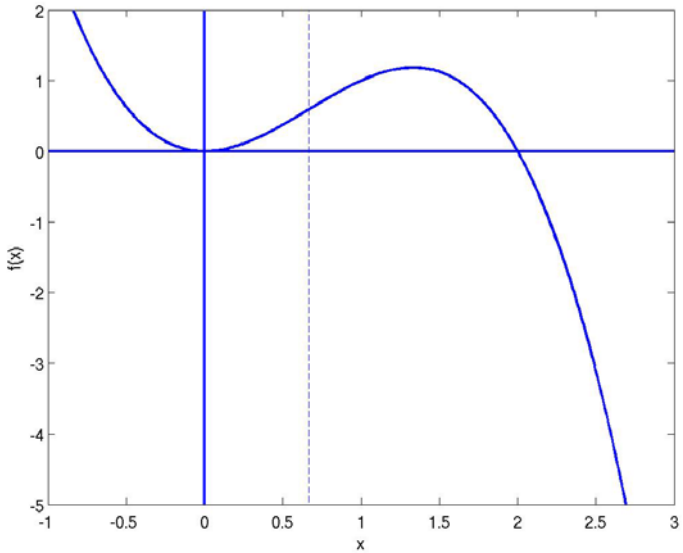
(3) $f(x) = 2x^2 - x^3$

$f'(x) = 4\,x - 3\,x^2$

$f''(x) = 4 - 6\,x$

$x < \dfrac{2}{3}$ dann konvex

$x > \dfrac{2}{3}$ dann konkav



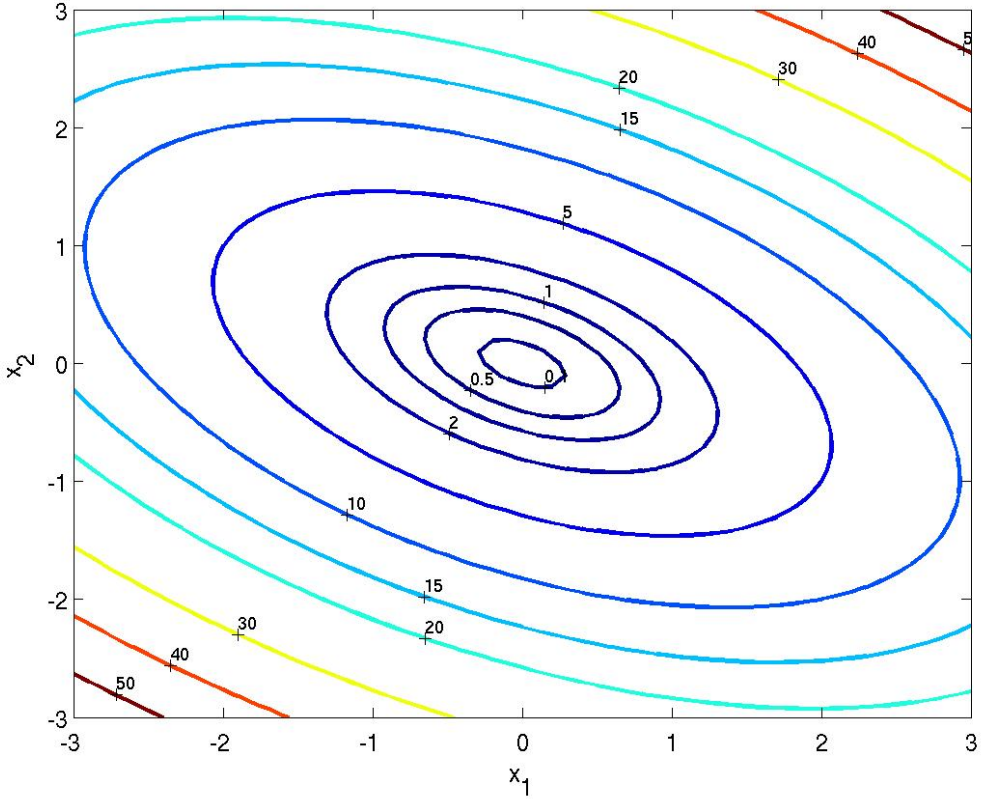
Beispiel: Konvex oder Konkav?

$$f(\mathbf{x}) = 3\,x_1^2 + 2\,x_1 x_2 + 1,5 x_2^2$$

$$\nabla f(\mathbf{x}) = \begin{bmatrix} \dfrac{\partial f}{\partial x_1} \\ \dfrac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 6\,x_1 + 2\,x_2 \\ 2\,x_1 + 3\,x_2 \end{bmatrix} = \mathbf{0} \qquad \mathbf{H} = \begin{bmatrix} 6 & 2 \\ 2 & 3 \end{bmatrix}$$

$$6 > 0,\ det \begin{bmatrix} 6 & 2 \\ 2 & 3 \end{bmatrix} > 0,$$

$\mathbf{H}$ positiv definit,
$f(x)$ konvex



Beispiel: $f(\mathbf{x}) = 4 + 4,5\,x_1 - 4\,x_2 + x_1^2 + 2\,x_2^2 - 2\,x_1 x_2 + x_1^4 - 2\,x_1^2 x_2$

$$\nabla f(\mathbf{x}) = \begin{bmatrix} 4,5 + 2\,x_1 - 2\,x_2 + 4\,x_1^3 - 4\,x_1 x_2 \\ -4 + 4\,x_2 - 2\,x_1 - 2\,x_1^2 \end{bmatrix} = \mathbf{0}$$

$$\mathbf{H}(\mathbf{x}) = \begin{bmatrix} 2 + 12\,x_1^2 - 4\,x_2 & -2 - 4\,x_1 \\ -2 - 4\,x_1 & 4 \end{bmatrix}$$

$$\mathbf{x}^{(1)} = \begin{bmatrix} 1,941 \\ 3,854 \end{bmatrix}, \quad \mathbf{H}^{(1)} = \begin{bmatrix} 31,794 & -9,764 \\ -9,764 & 4 \end{bmatrix}$$

positiv definit (Minimumpunkt), $f^{(1)} = 0,9855;$

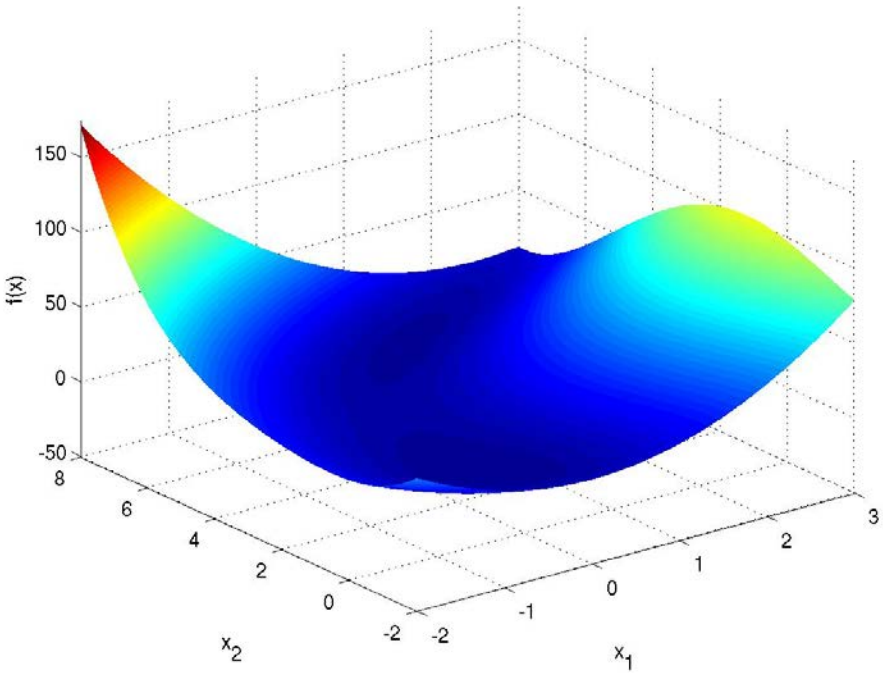
$$\mathbf{x}^{(2)} = \begin{bmatrix} -1,053 \\ 1,028 \end{bmatrix}, \quad \mathbf{H}^{(2)} = \begin{bmatrix} 11,194 & -6,212 \\ -6,212 & 4 \end{bmatrix}$$

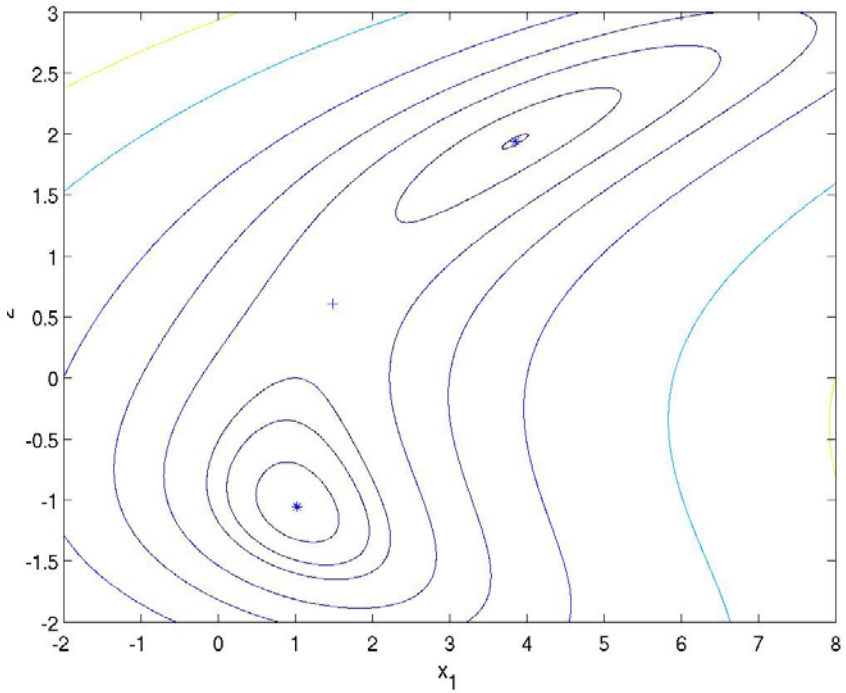
positiv definit (Minimumpunkt), $f^{(2)} = -0,5134;$

$$\mathbf{x}^{(3)} = \begin{bmatrix} 0,612 \\ 1,493 \end{bmatrix}, \quad \mathbf{H}^{(3)} = \begin{bmatrix} 0,523 & -4,448 \\ -4,448 & 4 \end{bmatrix}$$

indefinit (Sattelpunkt), $f^{(3)} = 2,83.$

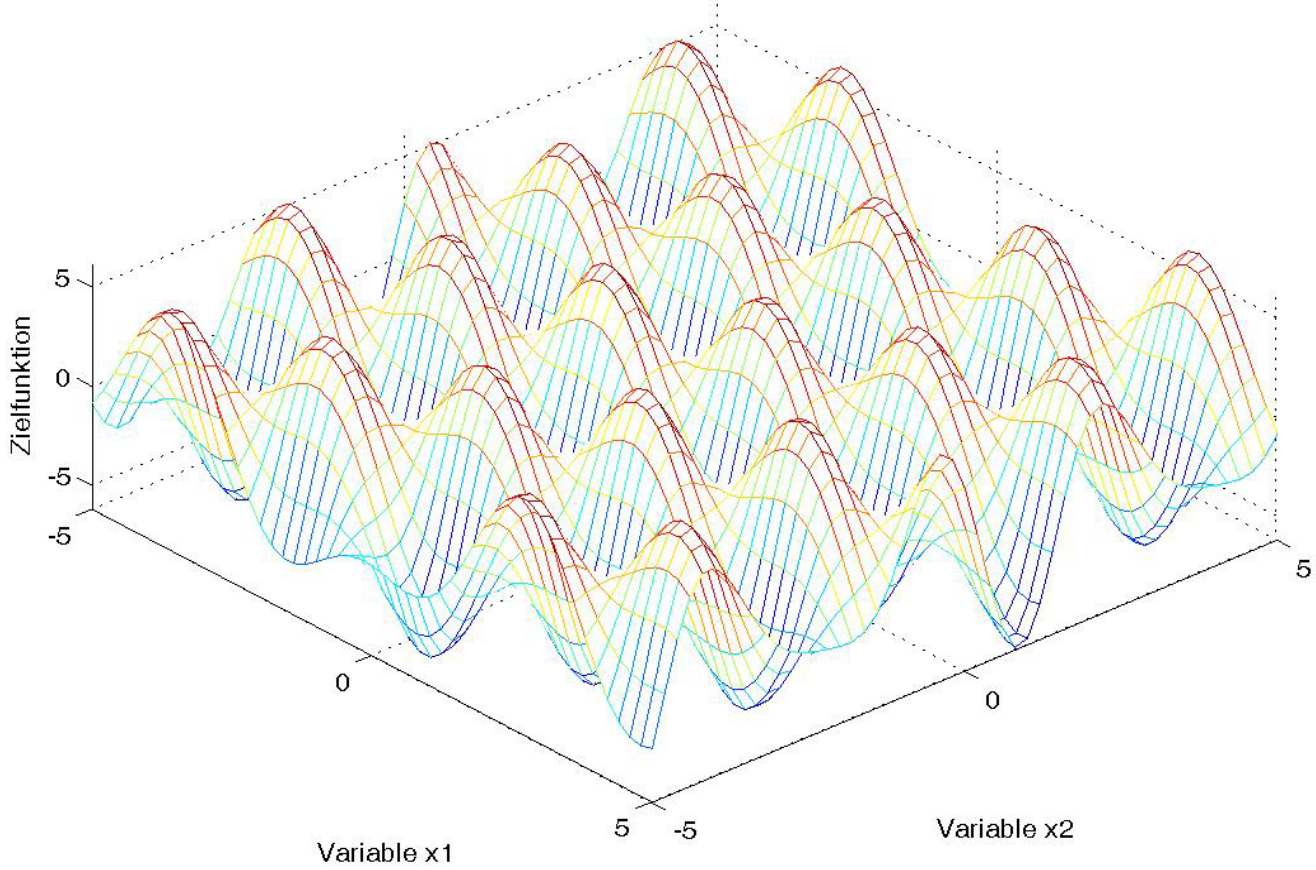
Beispiel: $f(\mathbf{x}) = 4 + 4,5\,x_1 - 4\,x_2 + x_1^2 + 2\,x_2^2 - 2\,x_1 x_2 + x_1^4 - 2\,x_1^2 x_2$





Beispiel:

$$f(x_1,\, x_2) = -0.2\sin(x_1 + 4x_2) - 2\cos(2x_1 + 3x_2) + 4\cos(x_1 - 2x_2)$$
$$-5 \leq x_1,\, x_2 \leq 5$$



Definition der Konvexität

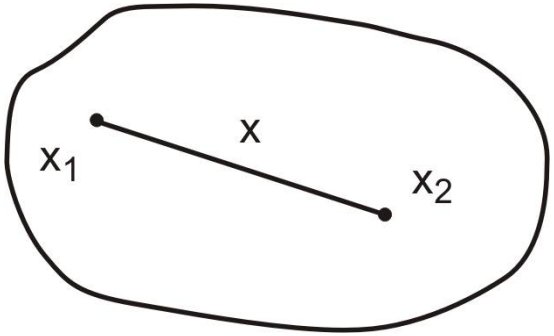
Eine konvexe Menge C

Zwei beliebige Punkte: $x_1,\ x_2 \subset C$ und $0 \leq \alpha \leq 1$

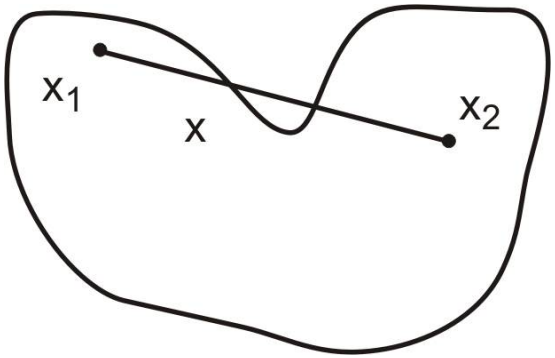
Es gibt: $x(\alpha) = \alpha x_1 + (1-\alpha)x_2$

und $x(\alpha) \subset C$

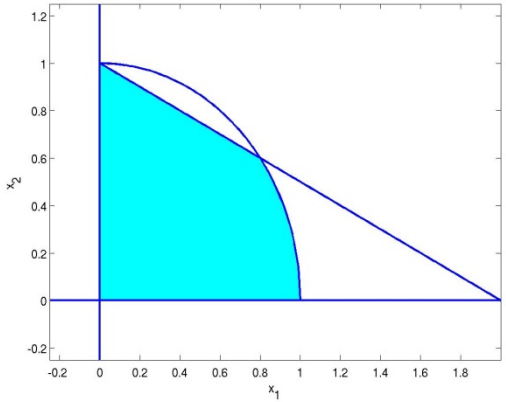
dann ist die Menge C konvex.



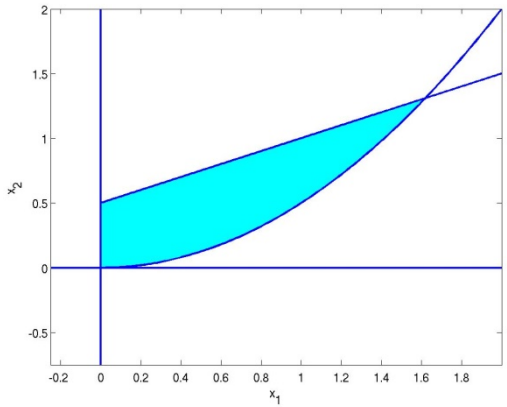
Sonst ist sie nichtkonvex.



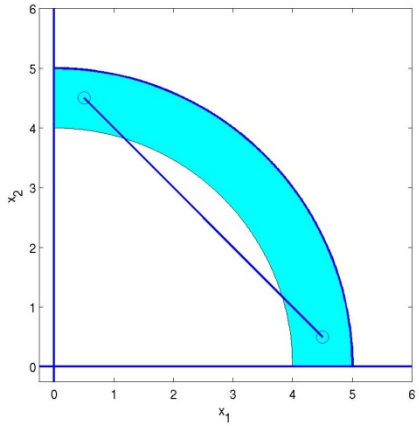
Konvexe und nichtkonvexe zulässige Bereiche



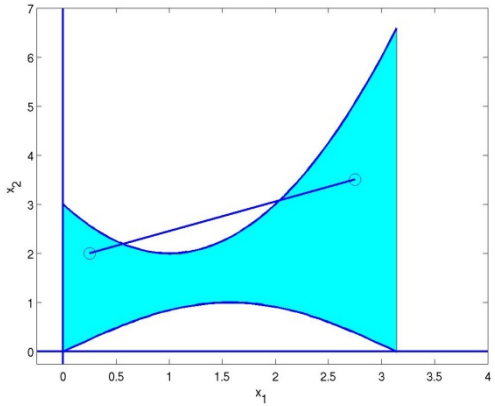
$$x_1^2 + x_2^2 \leq 1; \quad x_1 + 2x_2 \leq 2$$
$$x_1 \geq 0, \quad x_2 \geq 0$$



$$0,5\, x_1^2 - x_2 \leq 0; \quad x_1 - x_2 \geq -0,5$$
$$x_1 \geq 0, \quad x_2 \geq 0$$

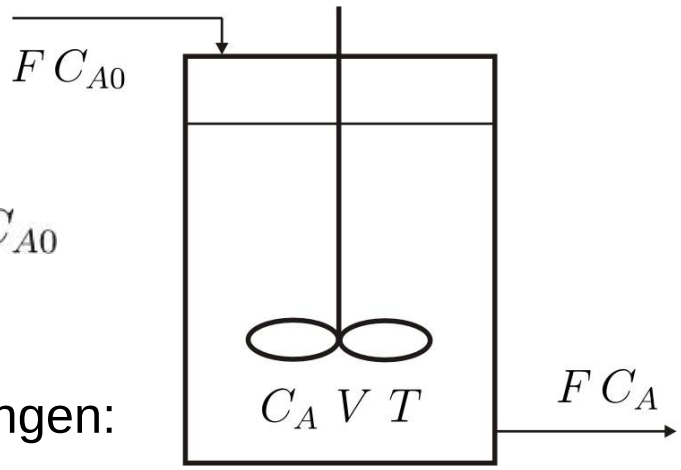


$$x_1^2 + x_2^2 \leq 25; \quad x_1^2 + x_2^2 \geq 16$$
$$x_1 \geq 0, \quad x_2 \geq 0$$



$$\sin(x_1) - x_2 \leq 0; \quad (x_1 - 1)^2 - x_2 + 2 \leq 0$$
$$3,14 \geq x_1 \geq 0, \quad x_2 \geq 0$$

Der zulässige Bereich beim Design eines Reaktors



Reaktionstyp: $A \rightarrow B$

Vorgegeben: Einsatzstrom F, Feedkonzentration C_{A0}

Anforderungen im Betrieb: Umsatz höher als 80 %

Entscheidungsvariablen: V, T mit den Beschränkungen:

$$0 \leq V \leq V_{max}$$

$$T_{min} \leq T \leq T_{max}$$

Reaktionsgeschwindigkeit: $r = kC_A$

Die Arrhenius-Gleichung: $k = k_0\, exp(-\frac{E}{RT})$

Komponentenbilanz: $F\,C_{A0} = F\,C_A + V\,k\,C_A$

Es folgt: $C_A = \frac{F\,C_{A0}}{F + V\,k}$

Definition des Umsatzes: $U_m = \dfrac{C_{A0} - C_A}{C_{A0}} \geq 0,8$

Es muss $U_m \geq 0,8$

Das bedeutet $C_{A0} - C_A \geq 0,8\, C_{A0}$

D. h. $C_A \leq 0,2\, C_{A0}$

Also $\dfrac{F\, C_{A0}}{F + V\, k} \leq 0,2\, C_{A0}$

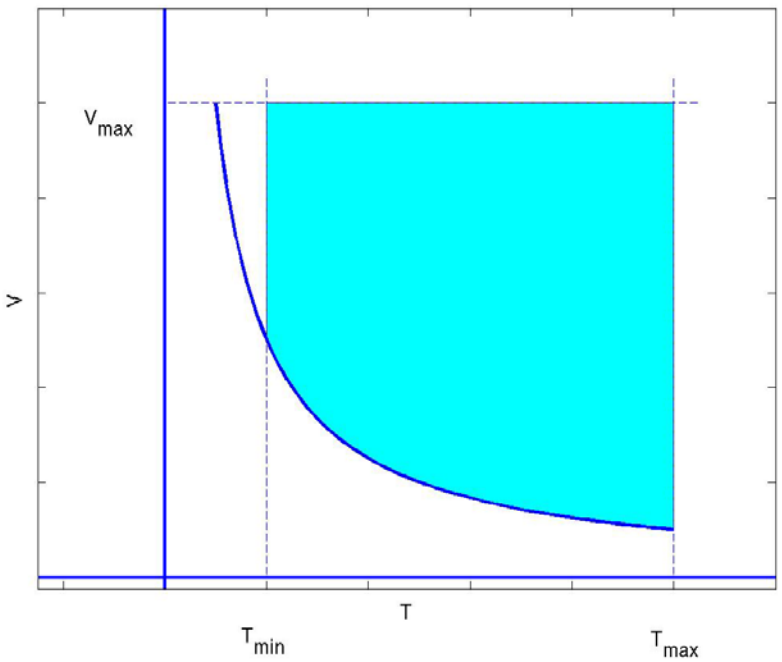
nämlich $F \leq 0,2(F + V\, k)$

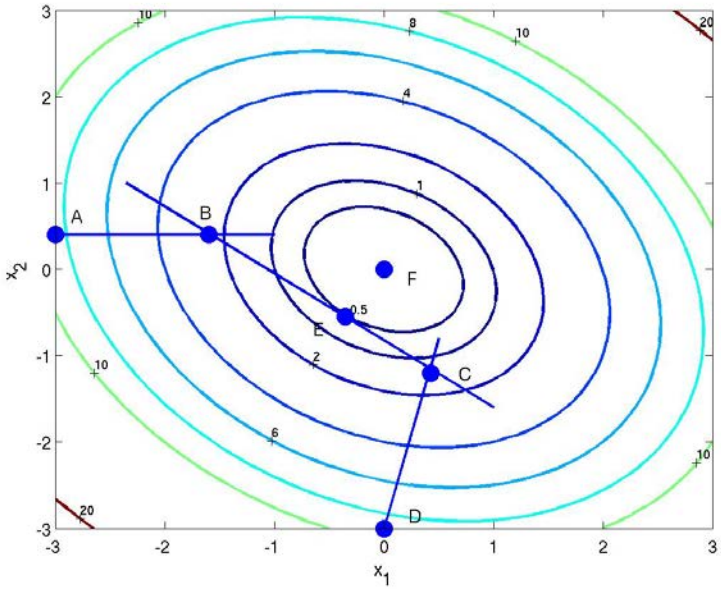
Es folgt $V\, k \geq 0,8\, F\,.$

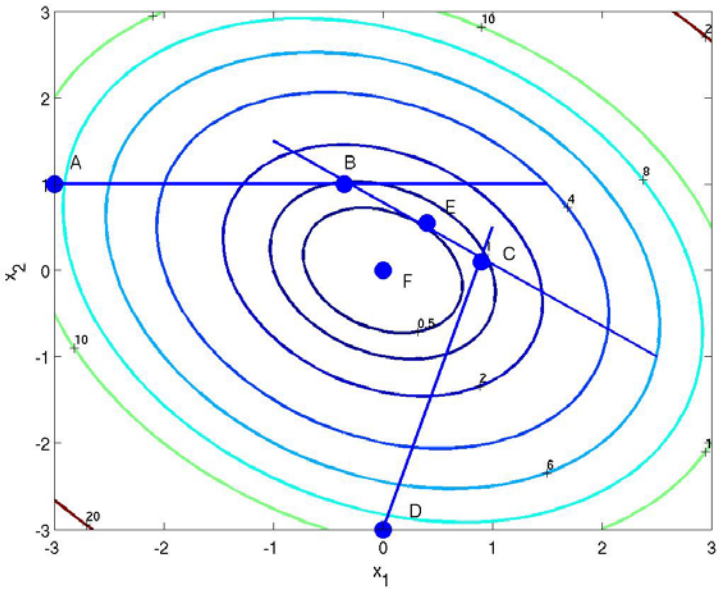
Anhand der Arrhenius-Gleichung für k,

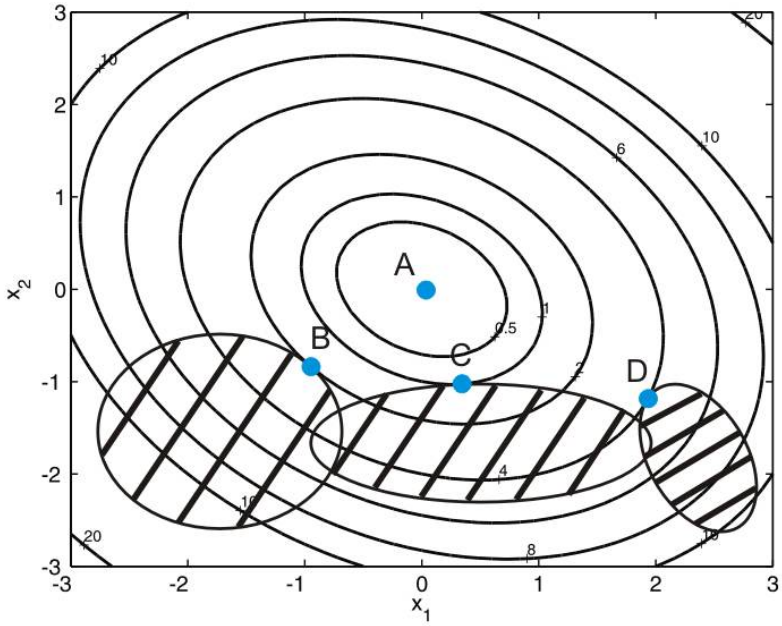
$$V \geq \frac{0,8\, F}{k_0}\, exp\left(\frac{E}{RT}\right)$$

Der zulässige Bereich für das Reaktordesign




A
B
C
D
E
F
x_1
x_2


A
B
C
D
E
F
x_1
x_2


A
B
C
D
x_1
x_2

Ein Optimierungsproblem ist ein konvexes Problem, wenn

- der zulässige Bereich eine konvexe Menge bildet und
- die Zielfunktion eine konvexe Funktion ist.

Zusammenfassung:

- Das lokale Optimum eines konvexen Problems ist das globale Optimum.
- Die Lösung eines nichtkonvexen Problems ist ein lokales Optimum und vom Schätzpunkt abhängig.
- Ein globaler Ansatz wird benötigt, um ein nichtkonvexes Problem zu lösen.

Ein Beispiel:

$$\min_{u,x} \quad \left\{(u-4)^2 + 1,1\,(x-2)^2\right\}$$

mit

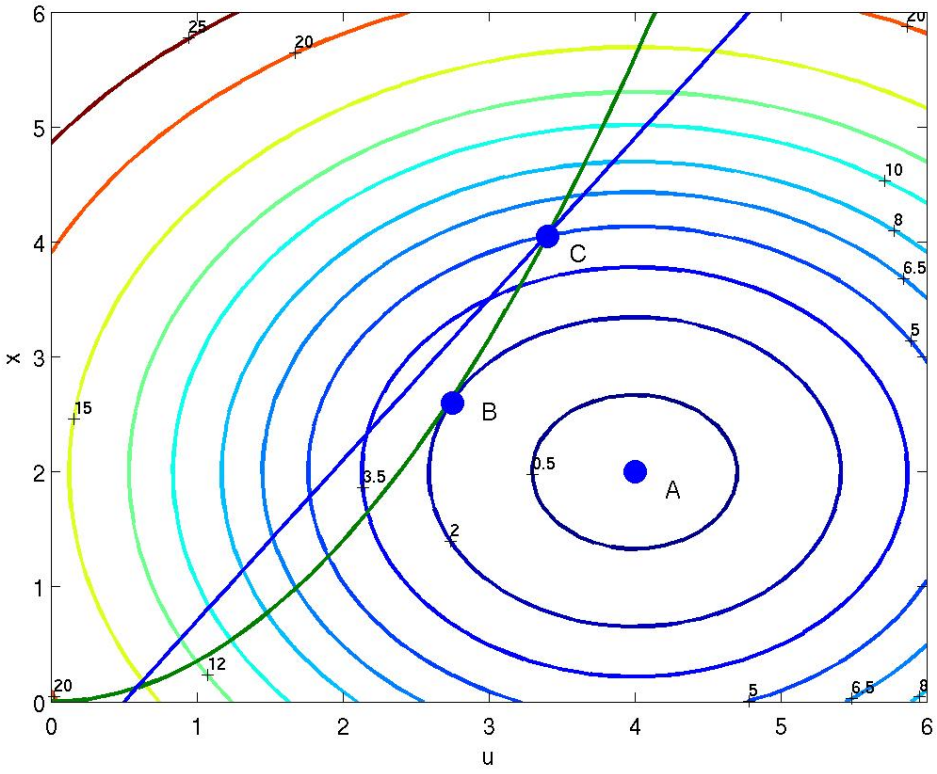
$$x - 0,35\,u^2 = 0$$

$$x + 0,7 - 1,4\,u \geq 0$$

$$0 \leq x \leq 6$$

$$0 \leq u \leq 6$$

Was bedeuten die Punkte A, B und C?



Fachgebiet Prozessoptimierung