Kapitel 11

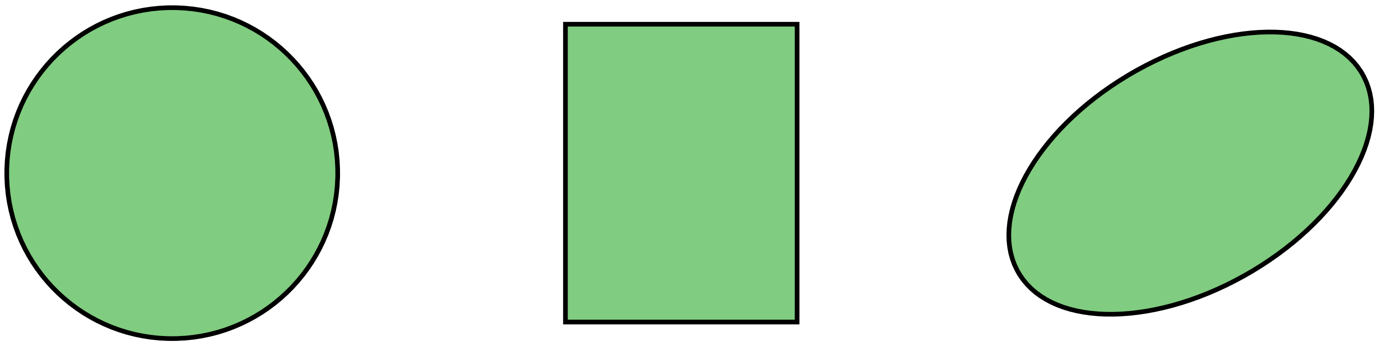
# Extrema

# Konvexe Menge

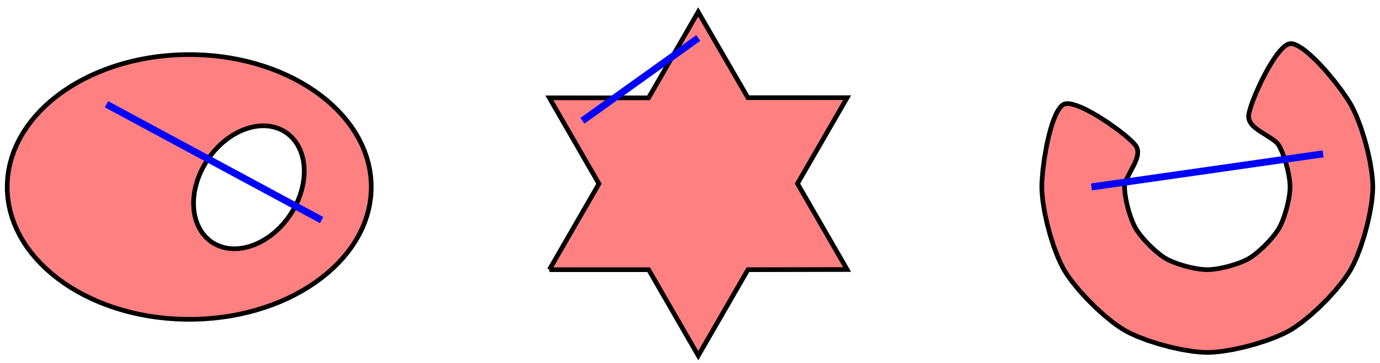
Eine Menge $D \subseteq \mathbb{R}^n$ heißt **konvex**, wenn für zwei beliebige Punkte $\mathbf{x}, \mathbf{y} \in D$ auch die Verbindungsstrecke dieser Punkte in $D$ liegt, d.h.

$$(1-h)\,\mathbf{x} + h\,\mathbf{y} \in D \qquad \text{für alle } h \in [0,1], \text{ und } \mathbf{x}, \mathbf{y} \in D$$

konvex:

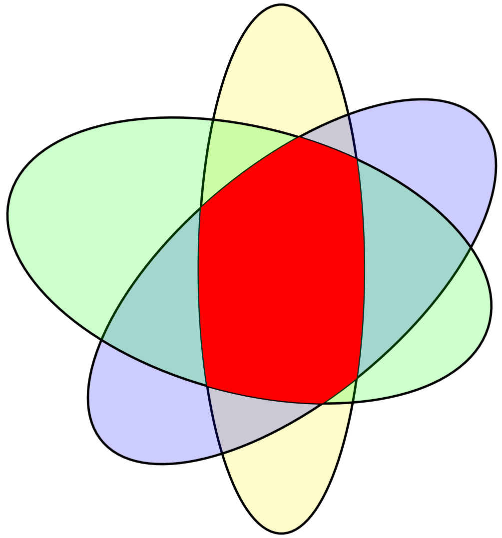


nicht konvex:



# Durchschnitt konvexer Mengen

Seien $S_1, \ldots, S_k$ konvexe Teilmengen des $\mathbb{R}^n$, dann ist auch der Durchschnitt $S_1 \cap \ldots \cap S_k$ konvex.



Die Vereinigung zweier konvexer Mengen muss nicht konvex sein.

# Beispiel – Halbräume

Seien $\mathbf{p} \in \mathbb{R}^n$ und $m$ fest, $\mathbf{p} \neq 0$. Dann beschreibt

$$H = \{\mathbf{x} \in \mathbb{R}^n \colon \mathbf{p}^t \cdot \mathbf{x} = m\}$$

eine **Hyperebene**, die den $\mathbb{R}^n$ in die zwei **Halbräume**

$$H_+ = \{\mathbf{x} \in \mathbb{R}^n \colon \mathbf{p}^t \cdot \mathbf{x} \geq m\}$$
$$H_- = \{\mathbf{x} \in \mathbb{R}^n \colon \mathbf{p}^t \cdot \mathbf{x} \leq m\}$$

teilt. Die Mengen $H$, $H_+$ und $H_-$ sind konvex.

Sei $\mathbf{x}$ ein Gütervektor, $\mathbf{p}$ der entsprechende Preisvektor und $m$ das verfügbare Einkommen. Dann ist die Budgetmenge konvex:

$$\{\mathbf{x} \in \mathbb{R}^n \colon \mathbf{p}^t \cdot \mathbf{x} \leq m, \mathbf{x} \geq 0\}$$
$$= \{\mathbf{x} \colon \mathbf{p}^t \cdot \mathbf{x} \leq m\} \cap \{\mathbf{x} \colon x_1 \geq 0\} \cap \ldots \cap \{\mathbf{x} \colon x_n \geq 0\}$$

# Konvexe und konkave Funktionen

Eine Funktion $f$ heißt **konvex** in $D \subseteq \mathbb{R}^n$, falls $D$ *konvex* ist, und

$$\boxed{f((1-h)\,\mathbf{x}_1 + h\,\mathbf{x}_2) \leq (1-h)\,f(\mathbf{x}_1) + h\,f(\mathbf{x}_2)}$$

für alle $\mathbf{x}_1, \mathbf{x}_2 \in D$ und alle $h \in [0,1]$.

Die Funktion $f$ heißt **konkav** in $D \subseteq \mathbb{R}^n$, falls $D$ *konvex* ist, und

$$\boxed{f((1-h)\,\mathbf{x}_1 + h\,\mathbf{x}_2) \geq (1-h)\,f(\mathbf{x}_1) + h\,f(\mathbf{x}_2)}$$

für alle $\mathbf{x}_1, \mathbf{x}_2 \in D$ und alle $h \in [0,1]$.

# Streng konvexe und streng konkave Funktionen

Eine Funktion $f$ heißt **streng konvex** in $D \subseteq \mathbb{R}^n$, falls $D$ *konvex* ist, und

$$\boxed{f((1-h)\,\mathbf{x}_1 + h\,\mathbf{x}_2) < (1-h)\,f(\mathbf{x}_1) + h\,f(\mathbf{x}_2)}$$

für alle $\mathbf{x}_1, \mathbf{x}_2 \in D$ mit $\mathbf{x}_1 \neq \mathbf{x}_2$ und alle $h \in (0,1)$.

Die Funktion $f$ heißt **streng konkav** in $D \subseteq \mathbb{R}^n$, falls $D$ *konvex* ist, und

$$\boxed{f((1-h)\,\mathbf{x}_1 + h\,\mathbf{x}_2) > (1-h)\,f(\mathbf{x}_1) + h\,f(\mathbf{x}_2)}$$

für alle $\mathbf{x}_1, \mathbf{x}_2 \in D$ mit $\mathbf{x}_1 \neq \mathbf{x}_2$ und alle $h \in (0,1)$.

# Beispiel – Lineare Funktion

Sei $\mathbf{a} \in \mathbb{R}^n$ konstant.
Dann ist $f(\mathbf{x}) = \mathbf{a}^t \cdot \mathbf{x}$ eine lineare Funktion und es gilt:

$$\begin{aligned} f((1-h)\,\mathbf{x}_1 + h\,\mathbf{x}_2) &= \mathbf{a}^t \cdot ((1-h)\,\mathbf{x}_1 + h\,\mathbf{x}_2) \\ &= (1-h)\,\mathbf{a}^t \cdot \mathbf{x}_1 + h\,\mathbf{a}^t \cdot \mathbf{x}_2 \\ &= (1-h)\,f(\mathbf{x}_1) + h\,f(\mathbf{x}_2) \end{aligned}$$

D.h., $f$ ist sowohl konkav als auch konvex.

Die lineare Funktion ist aber weder streng konkav noch streng konvex, da die Ungleichung niemals strikt ist.

# Beispiel – Quadratische Funktion in einer Variable

Die Funktion $f(x) = x^2$ ist streng konvex:

$$
\begin{aligned}
& f((1-h)\,x + h\,y) - \big[(1-h)\,f(x) + h\,f(y)\big] \\
&= ((1-h)\,x + h\,y)^2 - \big[(1-h)\,x^2 + h\,y^2\big] \\
&= (1-h)^2\,x^2 + 2(1-h)h\,xy + h^2\,y^2 - (1-h)\,x^2 - h\,y^2 \\
&= -h(1-h)\,x^2 + 2(1-h)h\,xy - h(1-h)\,y^2 \\
&= -h(1-h)\,(x-y)^2 \\
&< 0 \quad \text{für } x \neq y \text{ und } 0 < h < 1
\end{aligned}
$$

Also

$$f((1-h)\,x + h\,y) < (1-h)\,f(x) + h\,f(y)$$

für alle $x \neq y$ und $0 < h < 1$.

D.h. $f(x) = x^2$ ist streng konvex.

# Eigenschaften

- Falls $f(\mathbf{x})$ (streng) *konvex* ist, dann ist $-f(\mathbf{x})$ (streng) *konkav* (und umgekehrt).

- Falls $f_1(\mathbf{x}), \ldots, f_k(\mathbf{x})$ *konvex* (konkav) und $\alpha_1, \ldots, \alpha_k > 0$ ist, dann ist auch

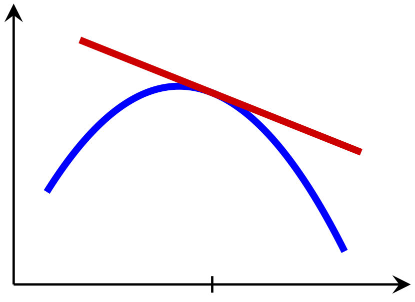
$$g(\mathbf{x}) = \alpha_1 f_1(\mathbf{x}) + \cdots + \alpha_k f_k(\mathbf{x})$$

  *konvex* (konkav).

- Falls (zumindest) eine der Funktionen $f_i(x)$ *streng konvex* (streng konkav) ist, dann ist auch $g(x)$ streng konvex (streng konkav).

# Eigenschaften

Für eine differenzierbare Funktion $f$ gilt:




- $f$ ist **konkav** genau dann, wenn

  $$\boxed{f(\mathbf{x}) - f(\mathbf{x}_0) \leq \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)}$$

  d.h., der Funktionsgraph liegt immer unterhalb der Tangente.

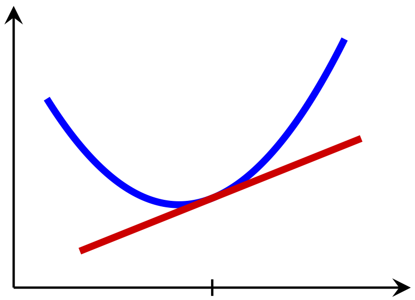
- $f$ ist **streng konkav** genau dann, wenn

  $$\boxed{f(\mathbf{x}) - f(\mathbf{x}_0) < \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) \qquad \text{für alle } \mathbf{x} \neq \mathbf{x}_0}$$

- $f$ ist **konvex** genau dann, wenn

  $$\boxed{f(\mathbf{x}) - f(\mathbf{x}_0) \geq \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)}$$

  (Analog für streng konvexe Funktion.)



# Quadratische Form

Sei $\mathbf{A}$ eine symmetrische Matrix und $q_{\mathbf{A}}(\mathbf{x}) = \mathbf{x}^t\mathbf{A}\mathbf{x}$ die entsprechende quadratische Form.

Wir können $\mathbf{A}$ diagonalisieren, d.h., es gibt eine Orthonormalbasis aus Eigenvektoren, sodass $\mathbf{A}$ zur Diagonalmatrix $\mathbf{D}$ aus Eigenwerten wird.

Wenn $\mathbf{c}$ der Koordinatenvektor von $\mathbf{x}$ bezüglich dieser Basis ist, dann ist

$$q_{\mathbf{A}}(\mathbf{x}) = c_1^2\lambda_1 + c_2^2\lambda_2 + \cdots + c_n^2\lambda_n$$

- Wenn alle Eigenwerte $\lambda_i \geq 0$ sind, dann ist $q_{\mathbf{A}}$ eine Summe von konvexen Funktionen und damit selbst konvex.
- Wenn alle$\lambda_i \leq 0$ sind, dann ist $q_{\mathbf{A}}$ konkav.
- Wenn es Eigenwerte $\lambda_i > 0$ und $\lambda_i < 0$, dann ist $q_{\mathbf{A}}$ weder konvex noch konkav.

# Quadratische Form

Für eine quadratische Form $q_{\mathbf{A}}$ gilt:

- *streng konvex* $\Leftrightarrow$ *positiv definit*
- *konvex* $\Leftrightarrow$ *positiv semidefinit*
- *streng konkav* $\Leftrightarrow$ *negativ definit*
- *konkav* $\Leftrightarrow$ *negativ semidefinit*
- weder noch $\Leftrightarrow$ *indefinit*

Die Definitheit von $\mathbf{A}$ kann festgestellt werden mit Hilfe

- der Eigenwerte von $\mathbf{A}$, oder
- der (allgemeinen) Hauptminoren von $\mathbf{A}$.

# Beispiel – Quadratische Form

Sei $\mathbf{A} = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 3 & -1 \\ 0 & -1 & 2 \end{pmatrix}$. Die Hauptminoren lauten:

$$H_1 = 2 \quad > 0$$

$$H_2 = \begin{vmatrix} 2 & 1 \\ 1 & 3 \end{vmatrix} = 5 \quad > 0$$

$$H_3 = |\mathbf{A}| = \begin{vmatrix} 2 & 1 & 0 \\ 1 & 3 & -1 \\ 0 & -1 & 2 \end{vmatrix} = 8 \quad > 0$$

$\mathbf{A}$ ist daher positiv definit.
Die quadratische Form $q_{\mathbf{A}}$ ist somit *streng konvex*.

# Beispiel – Quadratische Form

Sei $\mathbf{A} = \begin{pmatrix} -1 & 0 & 1 \\ 0 & -4 & 2 \\ 1 & 2 & -2 \end{pmatrix}$. Allgemeinen Hauptminoren:

$\tilde{H}_1 = -1$ $\quad$ $\tilde{H}_2 = -4$ $\quad$ $\tilde{H}_3 = -2$

$\tilde{H}_{1,2} = \begin{vmatrix} -1 & 0 \\ 0 & -4 \end{vmatrix} = 4$ $\quad$ $\tilde{H}_{1,3} = \begin{vmatrix} -1 & 1 \\ 1 & -2 \end{vmatrix} = 1$ $\quad$ $\tilde{H}_{2,3} = \begin{vmatrix} -4 & 2 \\ 2 & -2 \end{vmatrix} = 4$

$$\tilde{H}_{1,2,3} = \begin{vmatrix} -1 & 0 & 1 \\ 0 & -4 & 2 \\ 1 & 2 & -2 \end{vmatrix} = 0$$

$\tilde{H}_i \quad \leq 0$

$\tilde{H}_{i,j} \quad \geq 0$

$\tilde{H}_{1,2,3} \quad \leq 0$

$\mathbf{A}$ ist daher negativ semidefinit.
Die quadratische Form $q_{\mathbf{A}}$ ist somit *konkav* (aber nicht streng konkav).

# Krümmung differenzierbarer Funktionen

Sei $f\colon D \subseteq \mathbb{R}^n \to \mathbb{R}$ mit Taylorreihenentwicklung

$$f(\mathbf{x}_0 + \mathbf{h}) \approx f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot \mathbf{h} + \tfrac{1}{2}\,\mathbf{h}^t \cdot \mathbf{H}_f(\mathbf{x}_0) \cdot \mathbf{h}$$

Die *Hesse-Matrix* $\mathbf{H}_f(\mathbf{x}_0)$ bestimmt die *Krümmung* von $f$ in der Nähe des Entwicklungspunkts $\mathbf{x}_0$.

- $\mathbf{H}_f(\mathbf{x}_0)$ *positiv definit* $\Rightarrow$ $f$ *streng konvex* um $\mathbf{x}_0$
- $\mathbf{H}_f(\mathbf{x}_0)$ *negativ definit* $\Rightarrow$ $f$ *streng konkav* um $\mathbf{x}_0$

- $\mathbf{H}_f(\mathbf{x})$ *positiv semidefinit* für alle $\mathbf{x} \in D$ $\Leftrightarrow$ $f$ *konvex* in $D$
- $\mathbf{H}_f(\mathbf{x})$ *negativ semidefinit* für alle $\mathbf{x} \in D$ $\Leftrightarrow$ $f$ *konkav* in $D$

# Vorgangsweise – streng konvex

**1.** Berechne Hesse-Matrix

$$\mathbf{H}_f(\mathbf{x}) = \begin{pmatrix} f_{x_1x_1}(\mathbf{x}) & f_{x_1x_2}(\mathbf{x}) & \cdots & f_{x_1x_n}(\mathbf{x}) \\ f_{x_2x_1}(\mathbf{x}) & f_{x_2x_2}(\mathbf{x}) & \cdots & f_{x_2x_n}(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ f_{x_nx_1}(\mathbf{x}) & f_{x_nx_2}(\mathbf{x}) & \cdots & f_{x_nx_n}(\mathbf{x}) \end{pmatrix}$$

**2.** Berechne die (führenden) *Hauptminoren* $H_i$.

**3.** ▸ $f$ *streng konvex* $\Leftrightarrow$ alle $H_k > 0$ für (fast) **alle** $\mathbf{x} \in D$

▸ $f$ *streng konkav* $\Leftrightarrow$ alle $(-1)^k H_k > 0$ für (fast) **alle** $\mathbf{x} \in D$

[ $(-1)^k H_k > 0$ heißt: $H_1, H_3, \ldots < 0$ und $H_2, H_4, \ldots > 0$ ]

**4.** Andernfalls ist $f$ in $D$ *weder* streng konvex *noch* streng konkav.

# Vorgangsweise – konvex

**1.** Berechne Hesse-Matrix

$$\mathbf{H}_f(\mathbf{x}) = \begin{pmatrix} f_{x_1x_1}(\mathbf{x}) & f_{x_1x_2}(\mathbf{x}) & \cdots & f_{x_1x_n}(\mathbf{x}) \\ f_{x_2x_1}(\mathbf{x}) & f_{x_2x_2}(\mathbf{x}) & \cdots & f_{x_2x_n}(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ f_{x_nx_1}(\mathbf{x}) & f_{x_nx_2}(\mathbf{x}) & \cdots & f_{x_nx_n}(\mathbf{x}) \end{pmatrix}$$

**2.** Berechne die *allgemeinen Hauptminoren* $\tilde{H}_{i_1,\dots,i_k}$.

**3.**
- $f$ *konvex* $\Leftrightarrow$ alle $\tilde{H}_{i_1,\dots,i_k} \geq 0$ für **alle** $\mathbf{x} \in D$.
- $f$ *konkav* $\Leftrightarrow$ alle $(-1)^k \tilde{H}_{i_1,\dots,i_k} \geq 0$ für **alle** $\mathbf{x} \in D$.

**4.** Andernfalls ist $f$ in $D$ *weder* konvex *noch* konkav.

Offensichtlich ist jede *streng* konkave (konvexe) Funktion auch konkav (bzw. konvex).

# Beispiel

Ist die Funktion (streng) konkav oder konvex?

$$f(x,y) = x^4 + x^2 - 2\,x\,y + y^2$$

**1.** Hesse-Matrix: $\mathbf{H}_f(\mathbf{x}) = \begin{pmatrix} 12\,x^2 + 2 & -2 \\ -2 & 2 \end{pmatrix}$

**2.** Hauptminoren:

$H_1 = 12\,x^2 + 2 \quad \boxed{> 0}$

$H_2 = |\mathbf{H}_f(\mathbf{x})| = 24\,x^2 \quad \boxed{> 0}$ für alle $x \neq 0$.

**3.** Alle Hauptminoren $> 0$ für (fast) alle $\mathbf{x}$

$\Rightarrow$ $f$ ist *streng konvex*. (und damit auch konvex)

# Beispiel – Cobb-Douglas Funktion

Sei $f(x,y) = x^\alpha y^\beta$ mit $\alpha, \beta \geq 0$ und $\alpha + \beta \leq 1$,
und $D = \{(x,y) \colon x, y \geq 0\}$.

Hesse-Matrix an der Stelle $\mathbf{x}$:

$$\mathbf{H}_f(\mathbf{x}) = \begin{pmatrix} \alpha(\alpha-1)\, x^{\alpha-2} y^\beta & \alpha\beta\, x^{\alpha-1} y^{\beta-1} \\ \alpha\beta\, x^{\alpha-1} y^{\beta-1} & \beta(\beta-1)\, x^\alpha y^{\beta-2} \end{pmatrix}$$

Allgemeine Hauptminoren:

$$\tilde{H}_1 = \underbrace{\alpha}_{\geq 0} \underbrace{(\alpha-1)}_{\leq 0} \underbrace{x^{\alpha-2} y^\beta}_{\geq 0} \quad \leq 0$$

$$\tilde{H}_2 = \underbrace{\beta}_{\geq 0} \underbrace{(\beta-1)}_{\leq 0} \underbrace{x^\alpha y^{\beta-2}}_{\geq 0} \quad \leq 0$$

# Beispiel – Cobb-Douglas Funktion

$$
\begin{aligned}
\tilde{H}_{1,2} &= |\mathbf{H}_f(\mathbf{x})| \\
&= \alpha(\alpha-1)\,x^{\alpha-2}y^{\beta}\cdot\beta(\beta-1)\,x^{\alpha}y^{\beta-2} - (\alpha\beta\,x^{\alpha-1}y^{\beta-1})^2 \\
&= \alpha(\alpha-1)\,\beta(\beta-1)\,x^{2\alpha-2}y^{2\beta-2} - \alpha^2\beta^2\,x^{2\alpha-2}y^{2\beta-2} \\
&= \alpha\beta[(\alpha-1)(\beta-1)-\alpha\beta]x^{2\alpha-2}y^{2\beta-2} \\
&= \underbrace{\alpha\beta}_{\geq 0}\ \underbrace{(1-\alpha-\beta)}_{\geq 0}\ \underbrace{x^{2\alpha-2}y^{2\beta-2}}_{\geq 0} \qquad \geq 0
\end{aligned}
$$

$\tilde{H}_1 \leq 0$ und $\tilde{H}_2 \leq 0$, und $\tilde{H}_{1,2} \geq 0$ **für alle** $(x,y) \in D$.
$f(x,y)$ ist daher *konkav* in $D$.

Für $0 < \alpha, \beta < 1$ und $\alpha + \beta < 1$ gilt sogar:
$H_1 = \tilde{H}_1 < 0$ und $H_2 = |\mathbf{H}_f(\mathbf{x})| > 0$ **für** fast **alle** $(x,y) \in D$.
$f(x,y)$ ist dann sogar *streng konkav*.

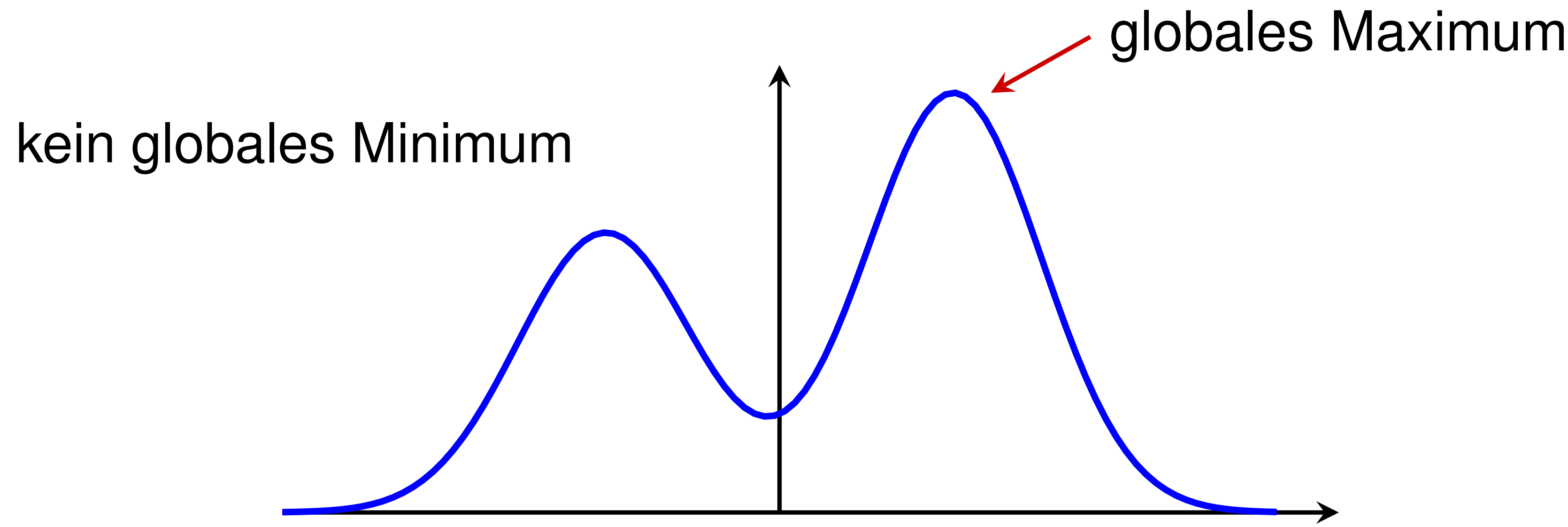
# Globales Extremum (Optimum)

Ein Punkt $\mathbf{x}^*$ heißt **globales Maximum** (*absolutes Maximum*) von $f$, falls für alle $\mathbf{x} \in D_f$ gilt:

$$f(\mathbf{x}^*) \geq f(\mathbf{x})$$

Ein Punkt $\mathbf{x}^*$ heißt **globales Minimum** (*absolutes Minimum*) von $f$, falls für alle $\mathbf{x} \in D_f$ gilt:

$$f(\mathbf{x}^*) \leq f(\mathbf{x})$$



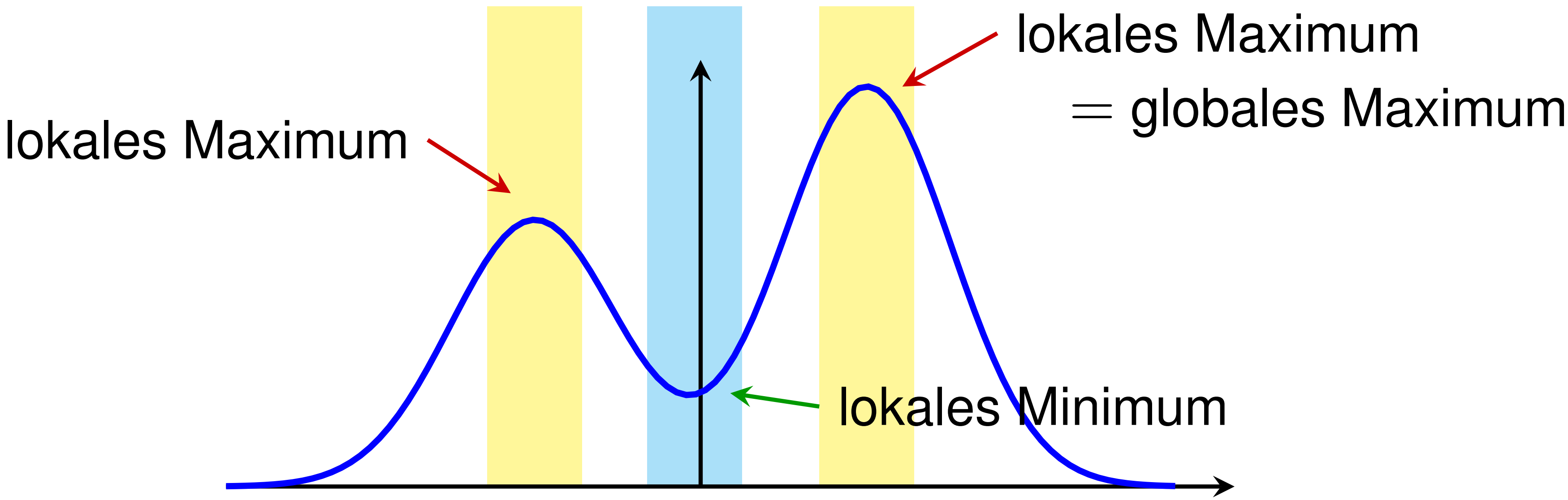
# Lokales Extremum (Optimum)

Ein Punkt $\mathbf{x}_0$ heißt **lokales Maximum** (*relatives* Maximum) von $f$, falls für alle $\mathbf{x}$ in einer geeigneten Umgebung von $\mathbf{x}_0$ gilt:

$$f(\mathbf{x}_0) \geq f(\mathbf{x})$$

Ein Punkt $\mathbf{x}_0$ heißt **lokales Minimum** (*relatives Minimum*) von $f$, falls für alle $\mathbf{x}$ in einer geeigneten Umgebung von $\mathbf{x}_0$ gilt:

$$f(\mathbf{x}_0) \leq f(\mathbf{x})$$



# Kritischer Punkt

In einem (lokalen) Maximum oder Minimum muss jede Richtungsableitung und damit auch der Gradient gleich Null sein.

Ein Punkt $\mathbf{x}_0$ heißt **kritischer Punkt** (oder *stationärer Punkt*) einer Funktion $f$, wenn

$$\nabla f(\mathbf{x}_0) = 0$$

*Notwendige Bedingung*:

Jedes Extremum von $f$ ist ein kritischer Punkt von $f$.

# Globales Extremum

*Hinreichende Bedingung*:

> Sei $\mathbf{x}_0$ ein kritischer Punkt einer **konkaven** (oder *konvexen*) Funktion $f$. Dann ist $\mathbf{x}_0$ ein **globales Maximum** (bzw. *globales Minimum*) von $f$.

Falls $f$ *streng* konkav (oder konvex) ist, dann ist das Extremum eindeutig bestimmt.

Die Aussage folgt unmittelbar aus den Eigenschaften (streng) konkaver Funktionen. Für alle $\mathbf{x} \neq \mathbf{x}_0$ gilt

$$f(\mathbf{x}) - f(\mathbf{x}_0) < \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) = 0 \cdot (\mathbf{x} - \mathbf{x}_0) = 0$$

und somit

$$f(\mathbf{x}_0) > f(\mathbf{x})$$

# Beispiel*

Sei $f(x) = e^x - 2\,x$.

Die Funktion ist streng konvex:

$$\begin{aligned} f'(x) &= e^x - 2 \\ f''(x) &= e^x \quad > 0 \quad \textbf{für alle } x \in \mathbb{R} \end{aligned}$$

Kritischer Punkt:

$$f'(x) = e^x - 2 = 0 \quad \Rightarrow \quad x_0 = \ln 2$$

$x_0 = \ln 2$ ist das (eindeutig bestimmte) globale Minimum von $f$.

# Beispiel

Sei $f\colon D = [0,\infty)^2 \to \mathbb{R},\ f(x,y) = 4\,x^{\frac{1}{4}}y^{\frac{1}{4}} - x - y$

Hesse-Matrix an der Stelle $\mathbf{x}$:

$$\mathbf{H}_f(\mathbf{x}) = \begin{pmatrix} -\frac{3}{4}\,x^{-\frac{7}{4}}y^{\frac{1}{4}} & \frac{1}{4}\,x^{-\frac{3}{4}}y^{-\frac{3}{4}} \\ \frac{1}{4}\,x^{-\frac{3}{4}}y^{-\frac{3}{4}} & -\frac{3}{4}\,x^{\frac{1}{4}}y^{-\frac{7}{4}} \end{pmatrix}$$

Hauptminoren:

$$H_1 = -\tfrac{3}{4}\,x^{-\frac{7}{4}}y^{\frac{1}{4}} \quad < 0$$

$$H_2 = \tfrac{1}{2}\,x^{-\frac{3}{2}}y^{-\frac{3}{2}} \quad > 0$$

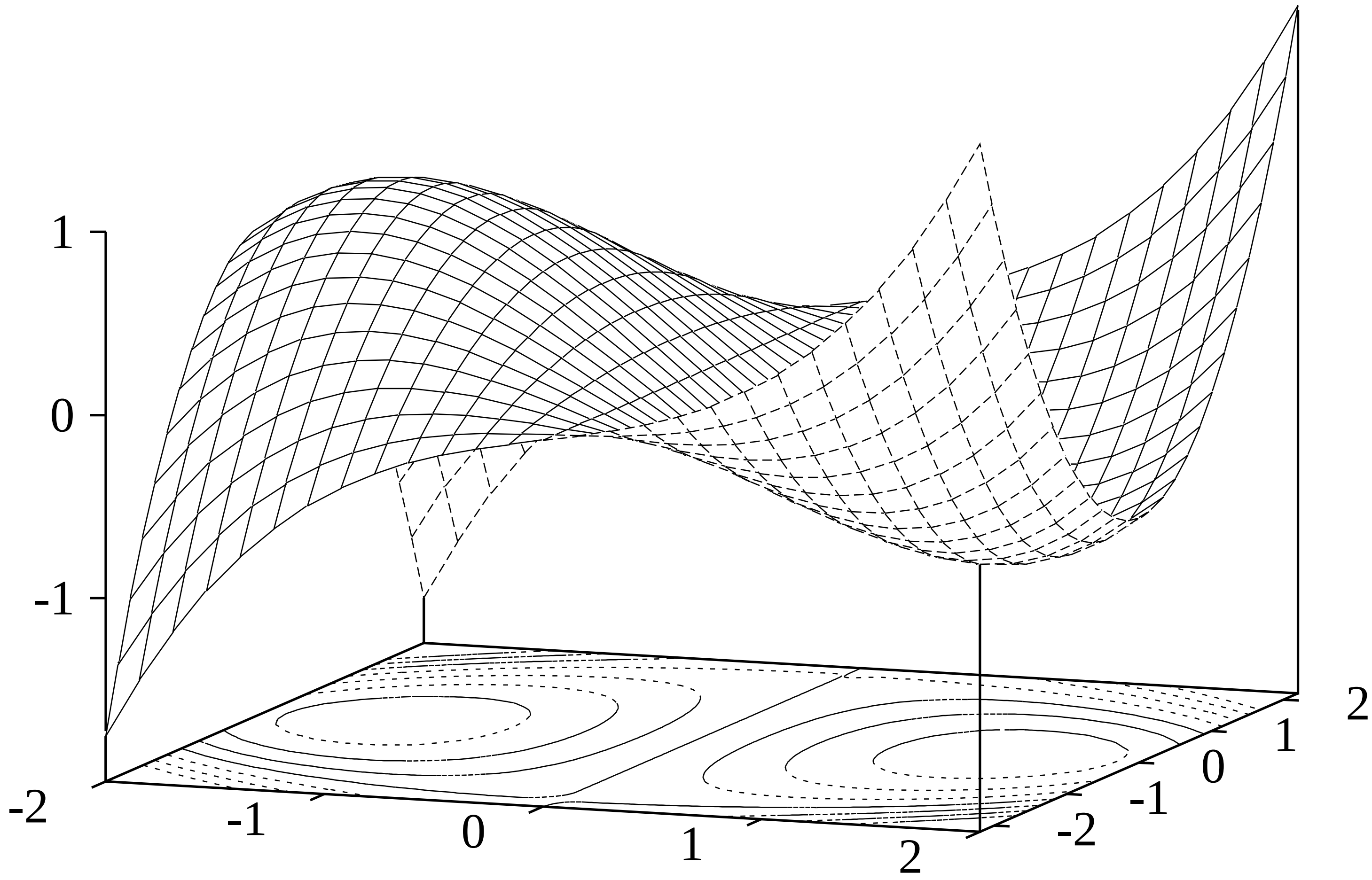
$f$ ist streng konkav in $D$.

Kritischer Punkt: $\nabla f = (x^{-\frac{3}{4}}y^{\frac{1}{4}} - 1, x^{\frac{1}{4}}y^{-\frac{3}{4}} - 1) = 0$

$$\begin{aligned} f_x &= x^{-\frac{3}{4}}y^{\frac{1}{4}} - 1 = 0 \\ f_y &= x^{\frac{1}{4}}y^{-\frac{3}{4}} - 1 = 0 \end{aligned} \qquad \Rightarrow \qquad \mathbf{x}_0 = (1,1)$$

$\mathbf{x}_0$ ist das globale Maximum von $f$.

# Beispiel – Lokale Extrema




$$f(x, y) = \frac{1}{6}\, x^3 - x + \frac{1}{4}\, x\, y^2$$

# Lokal konkav

Ein Punkt $\mathbf{x}_0$ ist ein **lokales Maximum** (oder *Minimum*) von $f$, falls

- $\mathbf{x}_0$ ist ein **kritischer Punkt** von $f$,
- $f$ ist **konkav** (bzw. *konvex*) um $\mathbf{x}_0$.

*Hinreichende Bedingung*:

$f$ ist *lokal* streng konkav (konvex) um $\mathbf{x}_0$ ist.

Sei $\mathbf{x}_0$ ein kritischer Punkt von $f$. Dann gilt

- $\mathbf{H}_f(\mathbf{x}_0)$ *negativ definit* $\Rightarrow$ $\mathbf{x}_0$ ist lokales Maximum
- $\mathbf{H}_f(\mathbf{x}_0)$ *positiv definit* $\Rightarrow$ $\mathbf{x}_0$ ist lokales Minimum

Es ist ausreichend die Hesse-Matrix $\mathbf{H}_f(\mathbf{x})$ am kritischen Punkt $\mathbf{x}_0$ auszuwerten. (Im Gegensatz zur Bedingung für globale Extrema.)

# Vorgangsweise – Univariate Funktion*

*Hinreichende Bedingung*
für lokale Extremwerte einer Funktion in *einer* Variablen:

1. Berechne $f'(x)$ und $f''(x)$.
2. Suche alle Punkte $x_i$ mit $f'(x_i) = 0$ (kritischen Punkte).
3. Falls $f''(x_i) < 0 \quad \Rightarrow \quad x_i$ ist ein *lokales Maximum*.
   Falls $f''(x_i) > 0 \quad \Rightarrow \quad x_i$ ist ein *lokales Minimum*.
   Falls $f''(x_i) = 0 \quad \Rightarrow$ *keine* Aussage möglich!

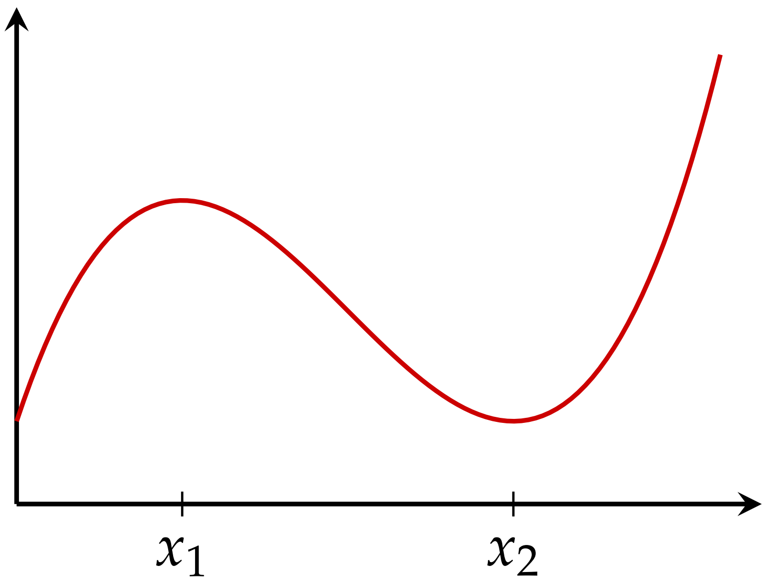
Falls $f''(x_i) = 0$ dann werden andere Methoden benötigt!

Z.B. kann man Terme höherer Ordnung der Taylorreihe mit Entwicklungspunkt $x_0$ betrachten.

# Beispiel – Univariate Funktion*

Gesucht sind die lokalen Extrema der Funktion

$$f(x) = \frac{1}{12}\,x^3 - x^2 + 3\,x + 1$$




1. $f'(x) = \frac{1}{4}\,x^2 - 2\,x + 3,$
   $f''(x) = \frac{1}{2}\,x - 2.$
2. $\frac{1}{4}\,x^2 - 2\,x + 3 = 0$
   besitzt die Lösungen
   $x_1 = 2$ und $x_2 = 6.$
3. $f''(2) = -1 \quad \Rightarrow \quad x_1$ ist lokales Maximum.
   $f''(6) = 1 \quad \Rightarrow \quad x_2$ ist lokales Minimum.

# Beispiel – Kritische Punkte

Suche alle kritischen Punkte von

$$f(x, y) = \frac{1}{6}\,x^3 - x + \frac{1}{4}\,x\,y^2$$

Partiellen Ableitungen:

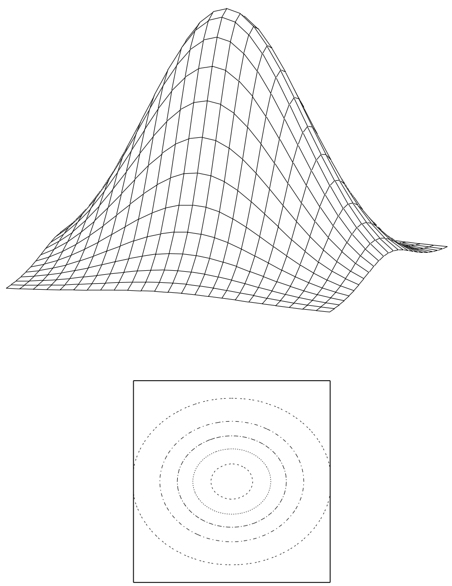
$$\begin{array}{lll} (I) & f_x = \tfrac{1}{2}\,x^2 - 1 + \tfrac{1}{4}\,y^2 & = 0 \\ (II) & f_y = \tfrac{1}{2}\,x\,y & = 0 \end{array}$$

$$\begin{array}{llcc|c} (II) & \Rightarrow & x = 0 & \text{oder} & y = 0 \\ (I) & \Rightarrow & -1 + \tfrac{1}{4}\,y^2 = 0 & & \tfrac{1}{2}\,x^2 - 1 = 0 \\ & & y = \pm 2 & & x = \pm\sqrt{2} \end{array}$$

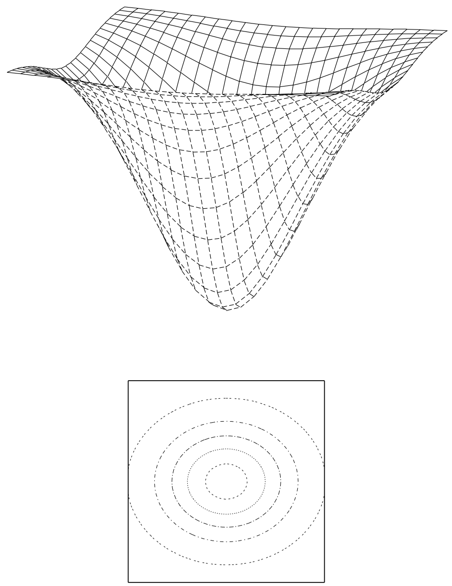
Kritische Punkte:

$$\begin{array}{ll} \mathbf{x}_1 = (0,\, 2) & \mathbf{x}_3 = (\sqrt{2},\, 0) \\ \mathbf{x}_2 = (0,\, -2) & \mathbf{x}_4 = (-\sqrt{2},\, 0) \end{array}$$

# Kritische Punkte – Lokale Extrema

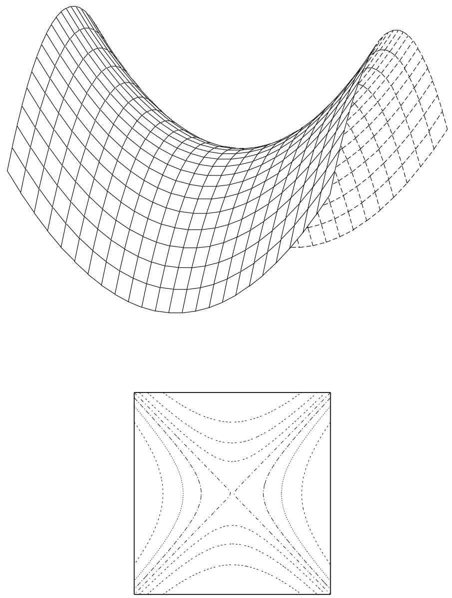


Lokales Maximum

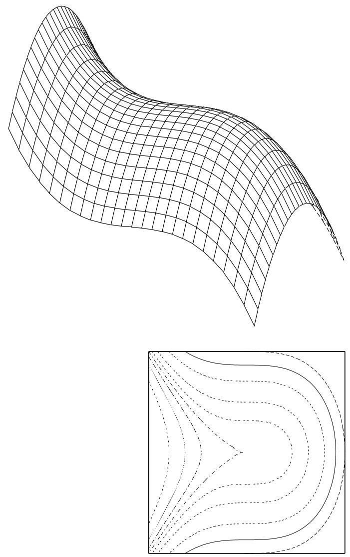


Lokales Minimum

# Kritische Punkte – Sattelpunkte



Sattelpunkt



Beispiel für höhere Ordnung

# Vorgangsweise – Lokale Extrema

1. Berechne Gradient $\nabla f$ und Hesse-Matrix $\mathbf{H}_f$.
2. Bestimme alle $\mathbf{x}_i$ mit $\nabla f(\mathbf{x}_i) = 0$ (kritische Punkte).
3. Berechne alle Hauptminoren $H_k$ für kritischen Punkt $\mathbf{x}_0$:

   **(a)** Alle Hauptminoren $H_k > 0$
   $\Rightarrow$ $\mathbf{x}_0$ ist ein **lokales Minimum** von $f$.

   **(b)** Für alle Hauptminoren gilt $(-1)^k H_k > 0$
   [ d.h., $H_1, H_3, \ldots < 0$ und $H_2, H_4, \ldots > 0$ ]
   $\Rightarrow$ $\mathbf{x}_0$ ist ein **lokales Maximum** von $f$.

   **(c)** $\det(\mathbf{H}_f(\mathbf{x}_0)) \neq 0$, aber weder (a) noch (b) sind erfüllt
   $\Rightarrow$ $\mathbf{x}_0$ ist ein **Sattelpunkt** von $f$.

   **(d)** Andernfalls ist *keine Aussage* möglich,
   d.h. $\mathbf{x}_0$ kann ein lokales Extremum sein, muss aber nicht.

# Vorgangsweise – Bivariate Funktion

1. Berechne Gradient $\nabla f$ und Hesse-Matrix $\mathbf{H}_f$.
2. Bestimme alle $\mathbf{x}_i$ mit $\nabla f(\mathbf{x}_i) = 0$ (kritische Punkte).
3. Berechne die Hauptminoren $H_1$ und $H_2$ für kritischen Punkt $\mathbf{x}_0$:

   **(a)** $H_2 > 0$ und $H_1 > 0$
   $\Rightarrow$ $\mathbf{x}_0$ ist ein **lokales Minimum** von $f$.

   **(b)** $H_2 > 0$ und $H_1 < 0$
   $\Rightarrow$ $\mathbf{x}_0$ ist ein **lokales Maximum** von $f$.

   **(c)** $H_2 < 0$
   $\Rightarrow$ $\mathbf{x}_0$ ist ein **Sattelpunkt** von $f$.

   **(d)** $H_2 = \det(\mathbf{H}_f(\mathbf{x}_0)) = 0$
   $\Rightarrow$ *keine Aussage* möglich.

# Beispiel – Bivariate Funktion

Suche die lokalen Extrema von

$$f(x,y) = \frac{1}{6}\,x^3 - x + \frac{1}{4}\,x\,y^2$$

**1.** $\nabla f = (\frac{1}{2}\,x^2 - 1 + \frac{1}{4}\,y^2, \frac{1}{2}\,x\,y)$

$$\mathbf{H}_f(x,y) = \begin{pmatrix} x & \frac{1}{2}\,y \\ \frac{1}{2}\,y & \frac{1}{2}\,x \end{pmatrix}$$

**2.** Kritische Punkte:

$$\mathbf{x}_1 = (0,2),\ \mathbf{x}_2 = (0,-2),\ \mathbf{x}_3 = (\sqrt{2},0),\ \mathbf{x}_4 = (-\sqrt{2},0)$$

# Beispiel – Bivariate Funktion (Forts.)

**3.** Hauptminoren:

$$\mathbf{H}_f(\mathbf{x}_1) = \mathbf{H}_f(0,2) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$H_2 = -1 \boxed{< 0} \quad \Rightarrow \quad \mathbf{x}_1 \text{ ist ein } \textit{Sattelpunkt}$$

$$\mathbf{H}_f(\mathbf{x}_2) = \mathbf{H}_f(0,-2) = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$$

$$H_2 = -1 \boxed{< 0} \quad \Rightarrow \quad \mathbf{x}_2 \text{ ist ein } \textit{Sattelpunkt}$$

# Beispiel – Bivariate Funktion (Forts.)

**3.** Hauptminoren:

$$\mathbf{H}_f(\mathbf{x}_3) = \mathbf{H}_f(\sqrt{2}, 0) = \begin{pmatrix} \sqrt{2} & 0 \\ 0 & \frac{\sqrt{2}}{2} \end{pmatrix}$$

$H_2 = 1 \; {>0}$ und $H_1 = \sqrt{2} \; {>0}$

$\Rightarrow \quad \mathbf{x}_3$ ist ein *lokales Minimum*

$$\mathbf{H}_f(\mathbf{x}_4) = \mathbf{H}_f(-\sqrt{2}, 0) = \begin{pmatrix} -\sqrt{2} & 0 \\ 0 & -\frac{\sqrt{2}}{2} \end{pmatrix}$$

$H_2 = 1 \; {>0}$ und $H_1 = -\sqrt{2} \; {<0}$

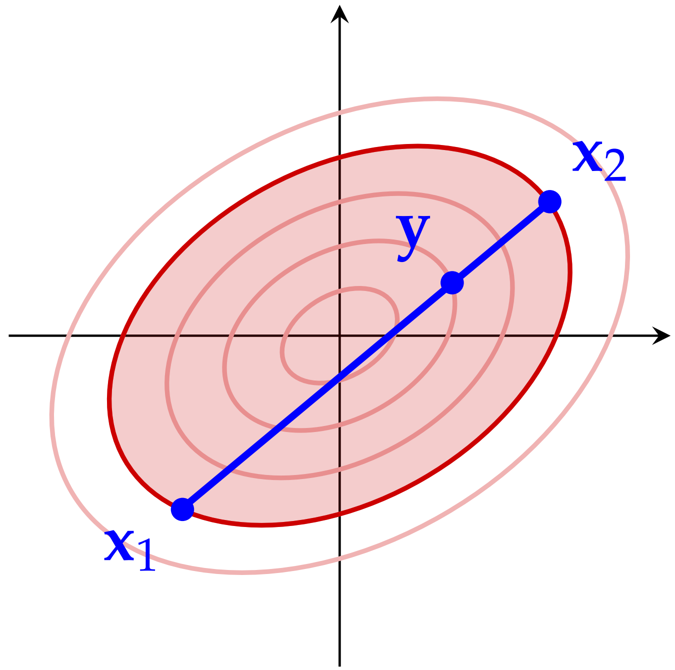
$\Rightarrow \quad \mathbf{x}_4$ ist ein *lokales Maximum*

# Konvexe Funktion und konvexe Menge

Sei $f$ *konvex*.
Dann ist die **untere Niveaumenge** von $f$

$$\{\mathbf{x} \in D_f : f(\mathbf{x}) \leq c\}$$

*konvex*.




Seien $\mathbf{x}_1, \mathbf{x}_2 \in \{\mathbf{x} \in D_f : f(\mathbf{x}) \leq c\}$, d.h. $f(\mathbf{x}_1), f(\mathbf{x}_2) \leq c$. Sei $\mathbf{y} = (1-h)\mathbf{x}_1 + h\mathbf{x}_2$ für ein $h \in [0,1]$.

Dann ist

$$\begin{aligned} f(\mathbf{y}) &= f((1-h)\mathbf{x}_1 + h\mathbf{x}_2) \\ &\leq (1-h)\,f(\mathbf{x}_1) + h\,f(\mathbf{x}_2) \\ &\leq (1-h)c + hc = c \end{aligned}$$

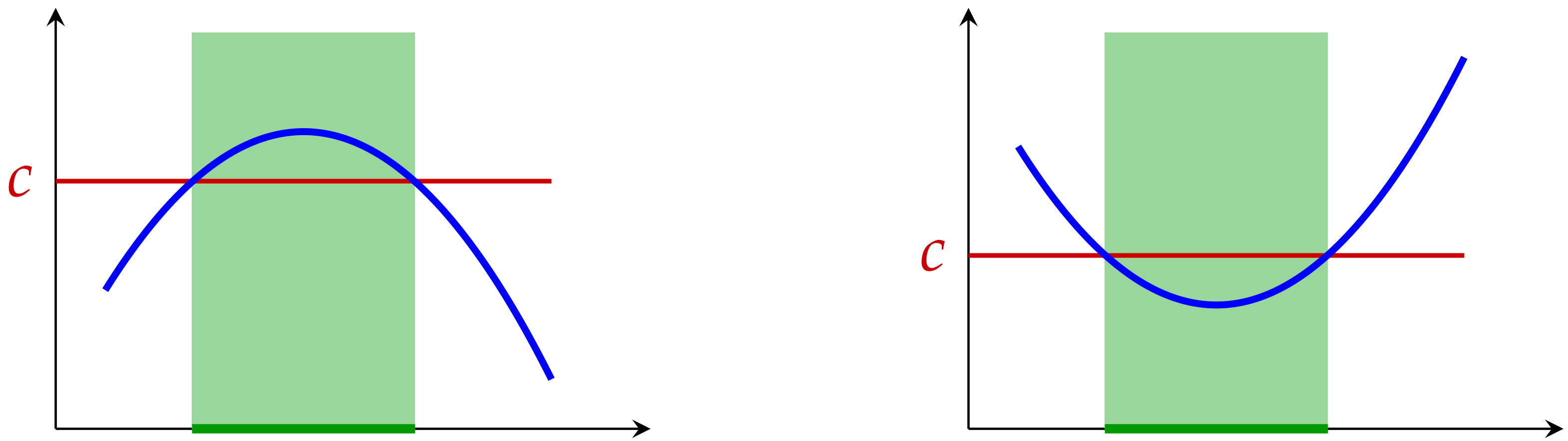
Also $\mathbf{y} \in \{\mathbf{x} \in D_f : f(\mathbf{x}) \leq c\}$.

# Konkave Funktion und konvexe Menge

Sei $f$ *konkav*.
Dann ist die **obere Niveaumenge** von $f$

$$\{\mathbf{x} \in D_f : f(\mathbf{x}) \geq c\}$$

*konvex*.




obere Niveaumenge　　untere Niveaumenge

# Extremum und monotone Transformation

Sei $T\colon \mathbb{R} \to \mathbb{R}$ eine *streng monoton steigende* Funktion.

Falls $\mathbf{x}^*$ ein *Maximum* (Minimum) von $f$ ist,
dann ist $\mathbf{x}^*$ auch ein Maximum (bzw. Minimum) von $T \circ f$:

Da $\mathbf{x}^*$ ein *Maximum* von $f$ ist, gilt

$$f(\mathbf{x}^*) \geq f(\mathbf{x}) \text{ für alle } \mathbf{x}.$$

Da $T$ streng monoton steigend ist, gilt

$$T(x_1) > T(x_2) \text{ falls } x_1 > x_2.$$

Wir erhalten daher

$$(T \circ f)(\mathbf{x}^*) = T(f(\mathbf{x}^*)) > T(f(\mathbf{x})) = (T \circ f)(\mathbf{x}) \text{ für alle } \mathbf{x},$$

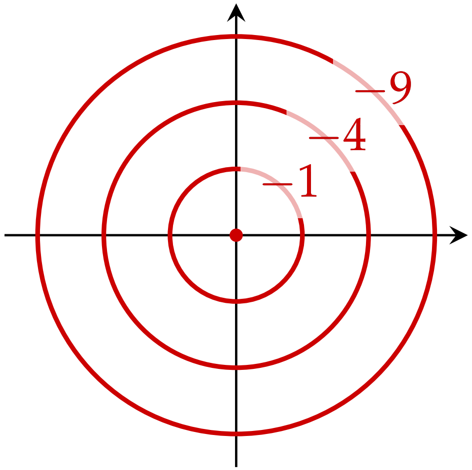
i.e., $\mathbf{x}^*$ ist ein Maximum von $T \circ f$.

Da $T$ injektiv ist, gilt sogar die Umkehrung:
Falls $\mathbf{x}^*$ ein *Maximum* (Minimum) von $T \circ f$ ist, dann auch von $f$.

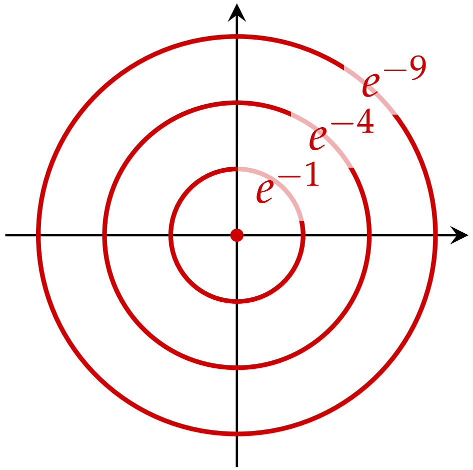
# Extremum und monotone Transformation

Die streng monotone Transformation $T$ erhält die Extrema von $f$.

$T$ erhält aber auch die Niveaumengen von $f$.




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




$$T(f(x,y)) = \exp(-x^2 - y^2)$$

# Quasi-konvex und quasi-konkav

Eine Funktion $f$ heißt **quasi-konvex** in $D \subseteq \mathbb{R}^n$, falls $D$ *konvex* ist und jede *untere Niveaumenge* $\{\mathbf{x} \in D_f : f(\mathbf{x}) \leq c\}$ *konvex* ist.

Eine Funktion $f$ heißt **quasi-konkav** in $D \subseteq \mathbb{R}^n$, falls $D$ *konvex* ist und jede *obere Niveaumenge* $\{\mathbf{x} \in D_f : f(\mathbf{x}) \geq c\}$ *konvex* ist.

# Konvex und quasi-konvex

Jede *konkave* (konvexe) Funktion ist auch *quasi-konkav* (bzw. quasi-konvex).

Eine quasi-konkave Funktion muss aber nicht konkav sein.

Sei $T$ eine streng monoton steigende Funktion. Falls eine Funktion $f(\mathbf{x})$ *konkav* (konvex) ist, dann ist $T \circ f$ *quasi-konkav* (bzw. quasi-konvex).

Die Funktion $g(x, y) = e^{-x^2-y^2}$ ist quasi-konkav, da $f(x, y) = -x^2 - y^2$ konkav ist, und $T(x) = e^x$ streng monoton steigend ist.
$g = T \circ f$ ist aber nicht konkav.

# Konvex, quasi-konvex und Extrema

Der Begriff *quasi-konvex* ist ein **schwächerer** Begriff als *konvex*, in dem Sinne, dass jede konvexe Funktion auch quasi-konvex ist, es aber viel mehr quasi-konvexe Funktionen gibt als konvexe.

Die Bedeutung dieses Begriffs liegt darin, dass sich manche Sätze über konvexe Funktionen auf quasi-konvexe Funktionen verallgemeinern lassen.

# Quasi-konvex und quasi-konkav II

- $f$ ist *quasi-konvex* genau dann, wenn

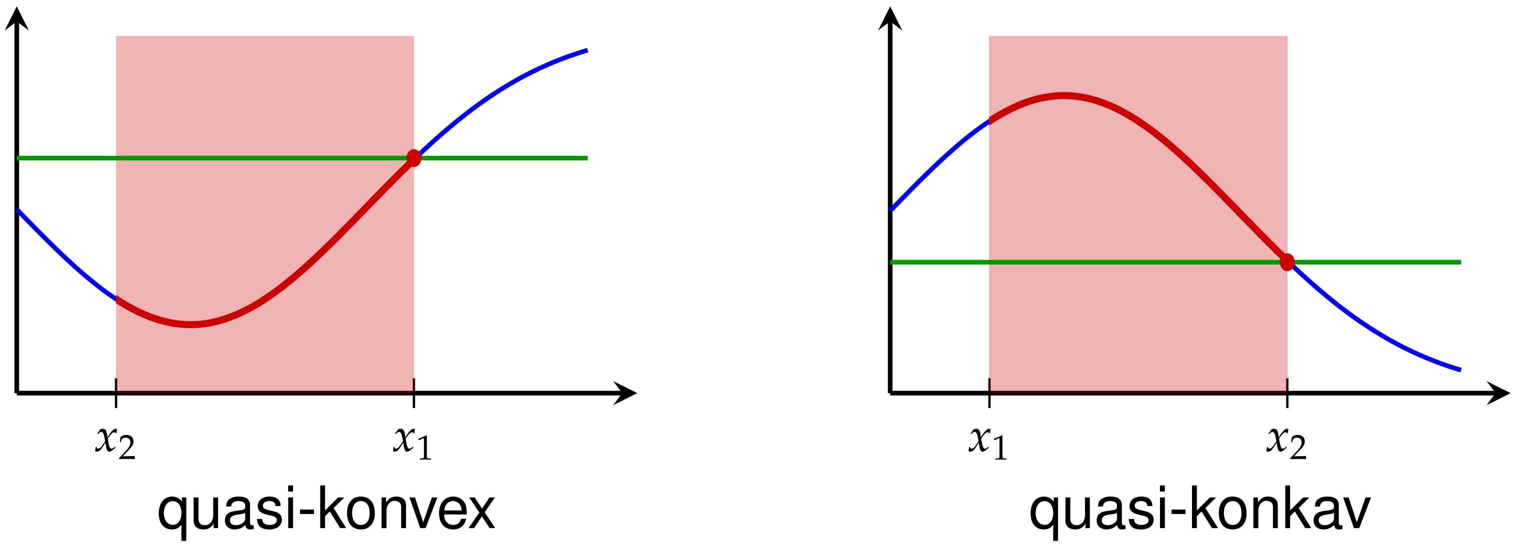
$$\boxed{f((1-h)\mathbf{x}_1 + h\mathbf{x}_2) \le \max\{f(\mathbf{x}_1), f(\mathbf{x}_2)\}}$$

für alle $\mathbf{x}_1, \mathbf{x}_2$ und $h \in [0,1]$.

- $f$ ist *quasi-konkav* genau dann, wenn

$$\boxed{f((1-h)\mathbf{x}_1 + h\mathbf{x}_2) \ge \min\{f(\mathbf{x}_1), f(\mathbf{x}_2)\}}$$

für alle $\mathbf{x}_1, \mathbf{x}_2$ und $h \in [0,1]$.



# Streng quasi-konvex und streng quasi-konkav

- Eine Funktion $f$ heißt **streng quasi-konvex**, wenn

  $$\boxed{f((1-h)\mathbf{x}_1 + h\mathbf{x}_2) < \max\{f(\mathbf{x}_1), f(\mathbf{x}_2)\}}$$

  für alle $\mathbf{x}_1, \mathbf{x}_2$, mit $\mathbf{x}_1 \neq \mathbf{x}_2$, und $h \in (0,1)$.

- Eine Funktion $f$ heißt **streng quasi-konkav**, wenn

  $$\boxed{f((1-h)\mathbf{x}_1 + h\mathbf{x}_2) > \min\{f(\mathbf{x}_1), f(\mathbf{x}_2)\}}$$

  für alle $\mathbf{x}_1, \mathbf{x}_2$, mit $\mathbf{x}_1 \neq \mathbf{x}_2$, und $h \in (0,1)$.

# Quasi-konvex und quasi-konkav III

Für eine differenzierbare Funktion $f$ gilt:

- $f$ ist *quasi-konvex* genau dann, wenn

$$\boxed{f(\mathbf{x}) \leq f(\mathbf{x}_0) \quad \Rightarrow \quad \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) \leq 0}$$

- $f$ ist *quasi-konkav* genau dann, wenn

$$\boxed{f(\mathbf{x}) \geq f(\mathbf{x}_0) \quad \Rightarrow \quad \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) \geq 0}$$

# Beispiel

Seien $p, r > 0$ und

$f\colon D = [0,\infty)^2 \to \mathbb{R},\quad f(x,y) = 4\,x^{\frac{1}{4}}y^{\frac{1}{4}} - px - ry$

Hesse-Matrix: $\mathbf{H}_f(\mathbf{x}) = \begin{pmatrix} -\frac{3}{4}\,x^{-\frac{7}{4}}y^{\frac{1}{4}} & \frac{1}{4}\,x^{-\frac{3}{4}}y^{-\frac{3}{4}} \\ \frac{1}{4}\,x^{-\frac{3}{4}}y^{-\frac{3}{4}} & -\frac{3}{4}\,x^{\frac{1}{4}}y^{-\frac{7}{4}} \end{pmatrix}$

Hauptminoren:

$H_1 = -\frac{3}{4}\,x^{-\frac{7}{4}}y^{\frac{1}{4}} \quad < 0$

$H_2 = \frac{1}{2}\,x^{-\frac{3}{2}}y^{-\frac{3}{2}} \quad > 0$

$f$ ist streng konkav in $D$.

Kritischer Punkt: $\nabla f = (x^{-\frac{3}{4}}y^{\frac{1}{4}} - p, x^{\frac{1}{4}}y^{-\frac{3}{4}} - r) = 0$

$f_x = x^{-\frac{3}{4}}y^{\frac{1}{4}} - p = 0$

$f_y = x^{\frac{1}{4}}y^{-\frac{3}{4}} - r = 0$

$\Rightarrow \quad \mathbf{x}_0 = \left(\sqrt{\frac{1}{r\,p^3}}, \sqrt{\frac{1}{r^3 p}}\right)$

$\mathbf{x}_0$ ist das globales Maximum von $f$.

**Frage:**
Wie ändert sich das Optimum $f^* = f(\mathbf{x}_0)$ mit den Parametern $r$ und $p$?

# Umhüllungssatz (Envelope-Theorem)

Gegeben sei eine Funktion

$$\mathbf{x} \mapsto f(\mathbf{x}, \mathbf{r}) \qquad \begin{array}{lll} \mathbf{x} = (x_1, \dots, x_n) \dots & \text{Variable} & \text{(endogen)} \\ \mathbf{r} = (r_1, \dots, r_k) \dots & \text{Parameter} & \text{(exogen)} \end{array}$$

mit Extremum $\mathbf{x}^*$.

Das Extremum hängt vom Parameter $\mathbf{r}$ ab:

$$\mathbf{x}^* = \mathbf{x}^*(\mathbf{r})$$

und damit auch der Extremalwert $f^*$:

$$f^*(\mathbf{r}) = f(\mathbf{x}^*(\mathbf{r}), \mathbf{r})$$

Es gilt:

$$\boxed{\frac{\partial f^*(\mathbf{r})}{\partial r_j} = \left.\frac{\partial f(\mathbf{x}, \mathbf{r})}{\partial r_j}\right|_{\mathbf{x}=\mathbf{x}^*(\mathbf{r})}}$$

# Umhüllungssatz (Envelope-Theorem)

$$
\begin{aligned}
\frac{\partial f^*(\mathbf{r})}{\partial r_j} &= \left.\frac{\partial f(\mathbf{x}^*(\mathbf{r}),\mathbf{r})}{\partial r_j}\right|_{\mathbf{x}=\mathbf{x}^*(\mathbf{r})} \qquad \text{[ Kettenregel ]} \\
&= \sum_{i=1}^{n} \underbrace{f_{x_i}(\mathbf{x}^*(\mathbf{r}),\mathbf{r})}_{\substack{=0 \\ \text{da kritischer Punkt}}} \cdot \frac{\partial x_i^*(\mathbf{r})}{\partial r_j} + \left.\frac{\partial f(\mathbf{x},\mathbf{r})}{\partial r_j}\right|_{\mathbf{x}=\mathbf{x}^*(\mathbf{r})} \\
&= \left.\frac{\partial f(\mathbf{x},\mathbf{r})}{\partial r_j}\right|_{\mathbf{x}=\mathbf{x}^*(\mathbf{r})}
\end{aligned}
$$

# Beispiel

Das (eindeutig bestimmte) Maximum von

$$f\colon D = [0,\infty)^2 \to \mathbb{R},\ \ f(x,y) = 4\,x^{\frac{1}{4}}y^{\frac{1}{4}} - px - ry$$

ist $\mathbf{x}^*(p,r) = (x^*(p,r), y^*(p,r)) = \left(\sqrt{\frac{1}{r\,p^3}}, \sqrt{\frac{1}{r^3 p}}\right)$.

**Frage:**
Wie ändert sich das Optimum $f^* = f(\mathbf{x}^*)$ mit den Parametern $r$ und $p$?

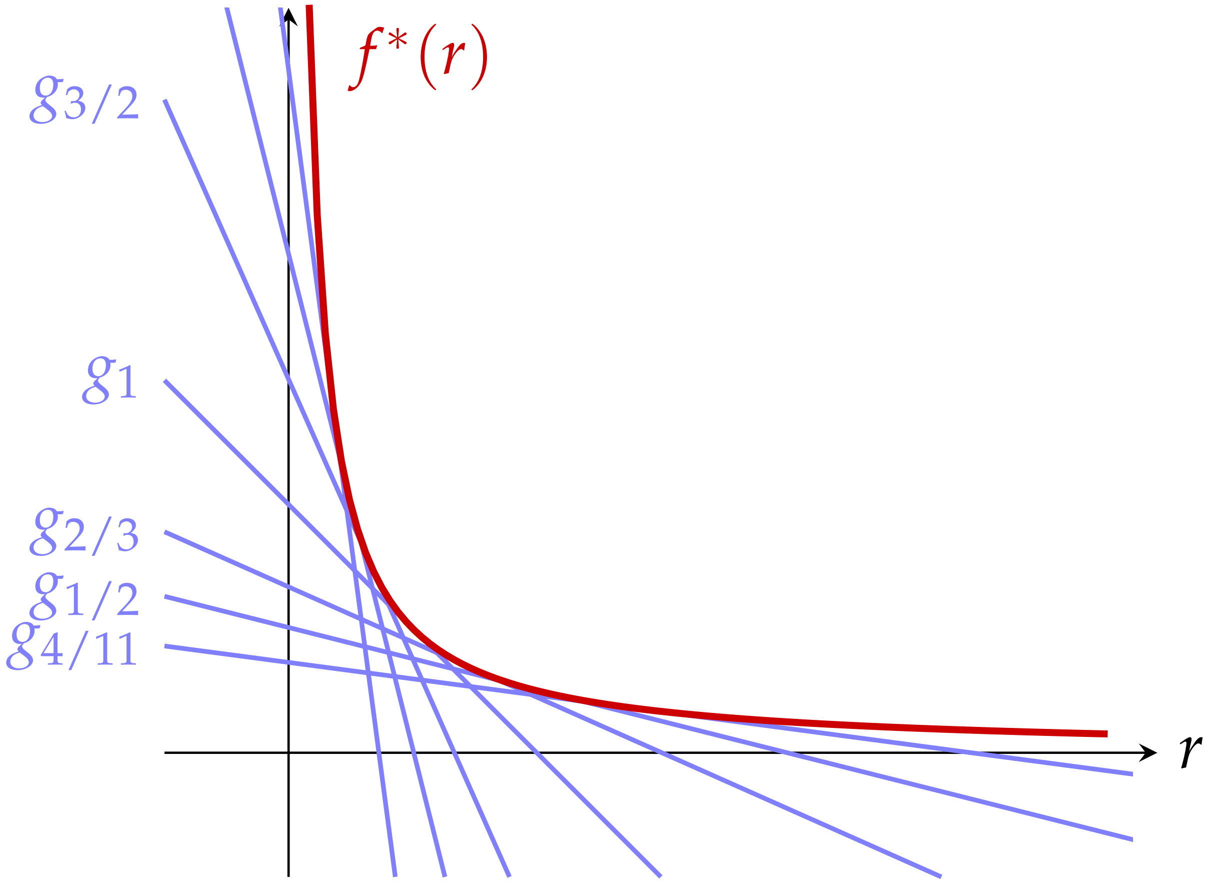
$$\frac{\partial f^*(p,r)}{\partial p} = \left.\frac{\partial f(\mathbf{x};p,r)}{\partial p}\right|_{\mathbf{x}=\mathbf{x}^*(p,r)} = -x\Big|_{\mathbf{x}=\mathbf{x}^*(p,r)} = -\sqrt{\frac{1}{r\,p^3}}$$

$$\frac{\partial f^*(p,r)}{\partial r} = \left.\frac{\partial f(\mathbf{x};p,r)}{\partial r}\right|_{\mathbf{x}=\mathbf{x}^*(p,r)} = -y\Big|_{\mathbf{x}=\mathbf{x}^*(p,r)} = -\sqrt{\frac{1}{r^3 p}}$$

# Eine geometrische Interpretation

Sei $f(x,r) = \sqrt{x} - rx$. Wir suchen $f^*(r) = \max_x f(x,r)$.

Zeichnen die Graphen von $g_x(r) = f(x,r)$ für verschiedene $x$.



# Zusammenfassung

- Konvexe Menge
- Konvex und konkav
- Quasi-konvex und quasi-konkav
- Lokales und globales Extremum
- Minimum, Maximum und Sattelpunkt
- Kritischer Punkt
- Hesse-Matrix und Hauptminor
- Umhüllungssatz (Envelope-Theorem)