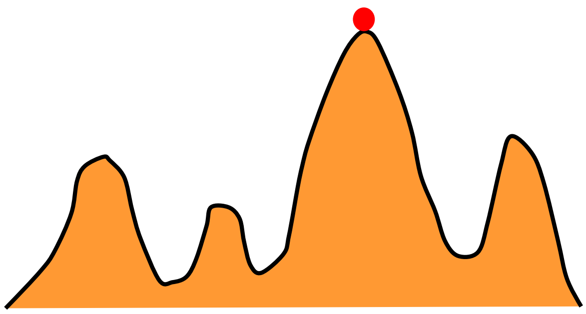
Kapitel 4: Optimierung

CHEMNITZ UNIVERSITY OF TECHNOLOGY

Prof. Dr. Fred Hamker

Department of Computer Science

Optimierung



Inhalt:

- Optimierung - anschaulich
- Bergsteigen
- Simuliertes ausglühen
- Genetische Algorithmen
- Vergleich: Suche und Optimieren
- Lokale Suche in kontinuierlichen Räumen
- Gradientenverfahren
- Lagrangefunktion

Empfohlene Literatur:

- Kapitel 4.3 und 4.4 in S. Russell und P. Norvig: Künstliche Intelligenz: Ein moderner Ansatz. München: Verlag Pearson Studium, 2005
- Kapitel 3 in B.-U. Köhler: Konzepte der statistischen Signalverarbeitung. Berlin, Heidelberg: Springer-Verlag, 2005.

Optimierung - anschaulich

Optimierungsverfahren sind ein unabdingbares Hilfsmittel für die Beantwortung vieler Fragestellungen

- Die Aufgabe besteht in der systematischen Suche von Extremwerten (von Funktionen).
- Die Elemente des Lösungsraumes sind Punkte in einer Landschaft.
- Die Evaluierungsfunktion ist die Höhe des Punktes
- Die Optimierung geschieht durch Bewegen in der Landschaft, bis man den höchsten Punkt erreicht hat.
- Es wird immer nur der aktuelle Punkt betrachtet (kein Gedächtnis, aber manchmal Momentum).

Bergsteigen

- einfache Schleife
- man versucht ***immer***, die aktuelle Lösung zu verbessern
- kein Suchbaum im Speicher, nur aktuelle Lösung (node) und deren Bewertung

```
function Bergsteigen(Problem) returns einen Lösungszustand
        inputs:   Problem ; eine Problembeschreibung
        local variables: node       ; der aktuelle Knoten
                         next       ; ein Knoten

        node ← Make-Node(Initial-Solution[Problem])
        loop do
        next ← ein Nachfolger von node mit höchstem Wert
                if Value[next] < Value[node] then return node (keine bessere
                                                        Lösung)
        node ← next
        end
```

Bergsteigen (Beispiel 8-Damen-Problem)

Schritt (Aktion): Wähle Dame aus und ändere ihre Position.

Mögliche Nachfolgezustände:

schiebe eine Dame (von 8) in ein anderes Quadrat derselben Spalte (8*7=56)

Bewertung (Minimum):

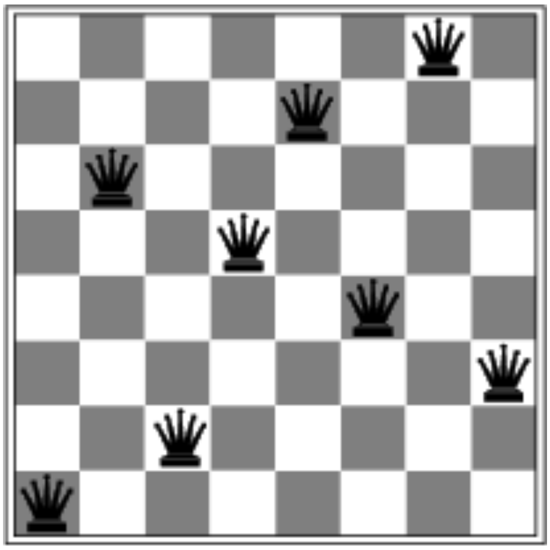
Heuristische Kostenfunktion h: Anzahl der Damenpaare, die einander direkt oder indirekt angreifen.

h(aktuell)=17

h(Nachfolger)=12

Bergsteigen (Beispiel 8-Damen-Problem)

5 Schritte

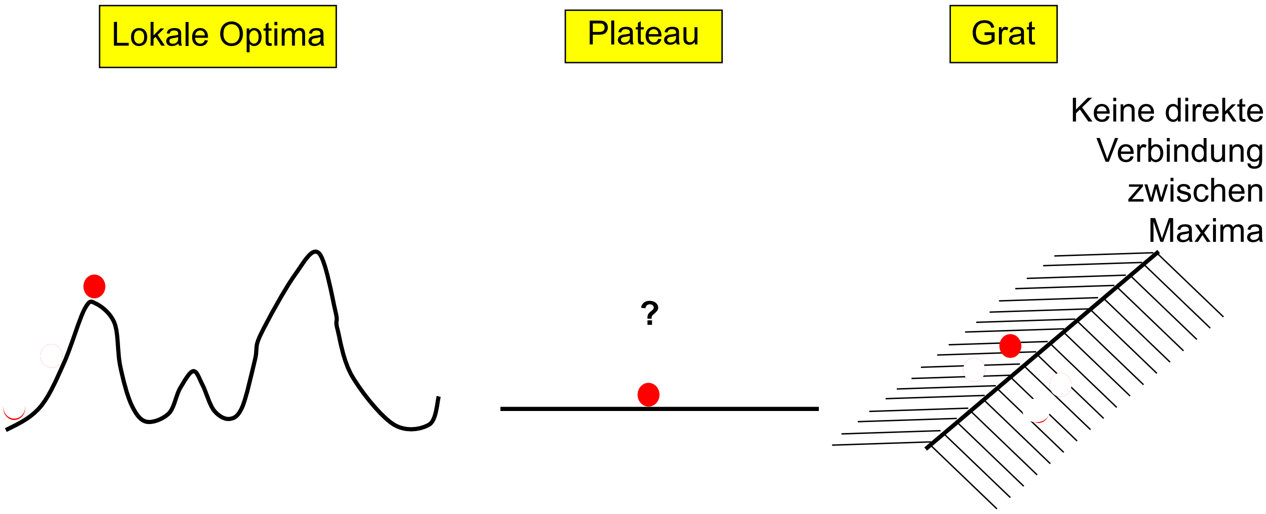


h=17

h=1 allerdings lokales Max.

Bei einem zufälligen Start bleibt das Bersteigen <u>mit steilem Anstieg</u> in 84% stecken und löst das Problem in 16%. Es funktioniert allerdings schnell (im Mittel 4 Schritte, wenn es das Problem löst, sonst 3) bei einem Zustandsraum von 8^8 (ca. 17 Millionen) Zuständen. Varianten: stochastisches Bergsteigen, Bergsteigen mit Neustart.

Bergsteigen - Probleme



Bersteigen ohne Abwärtszüge in Richtung von Zuständen mit geringerem Wert führt schnell zum steckenbleiben in lokalen Maxima. Die Wahl zufälliger Nachfolger ist allerdings ineffizient.

Potentielle Lösung: Erst mehr zufällig, später eher steil → Simulated Annealing

Simuliertes Ausglühen (Simulated Annealing)

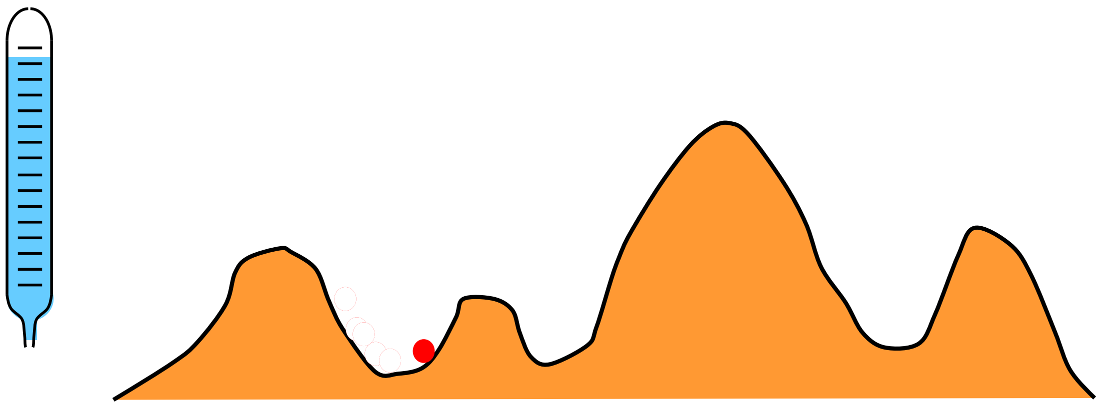
Idee: Während der Optimierung mit einer gewissen Wahrscheinlichkeit Zustände mit schlechteren Bewertungen auswählen. Die Wahrscheinlichkeit verringert sich mit der Zeit.

Wahrscheinlichkeit: → $\boxed{e^{\frac{\Delta E}{\mathrm{T}}}, \Delta E \leq 0}$

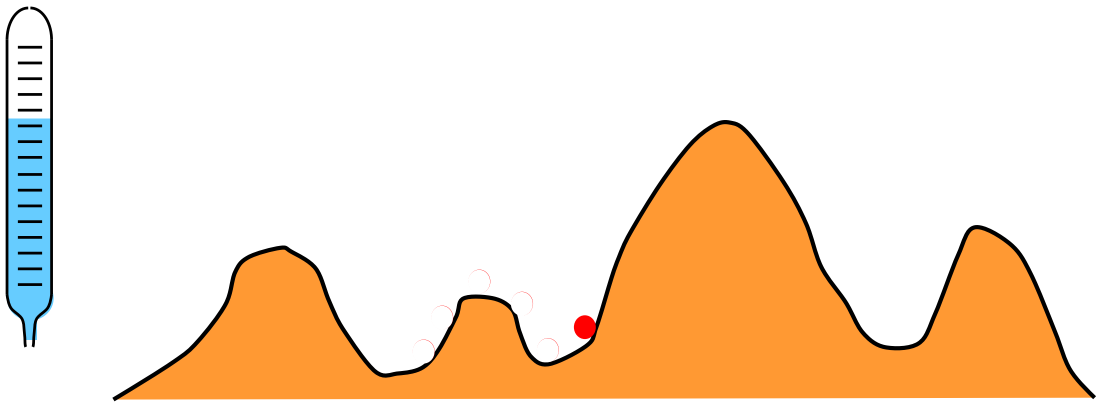
T (Temperatur) klein → Wahrscheinlichkeit wird klein

$T \to 0$ wird kleiner, je länger das Verfahren läuft

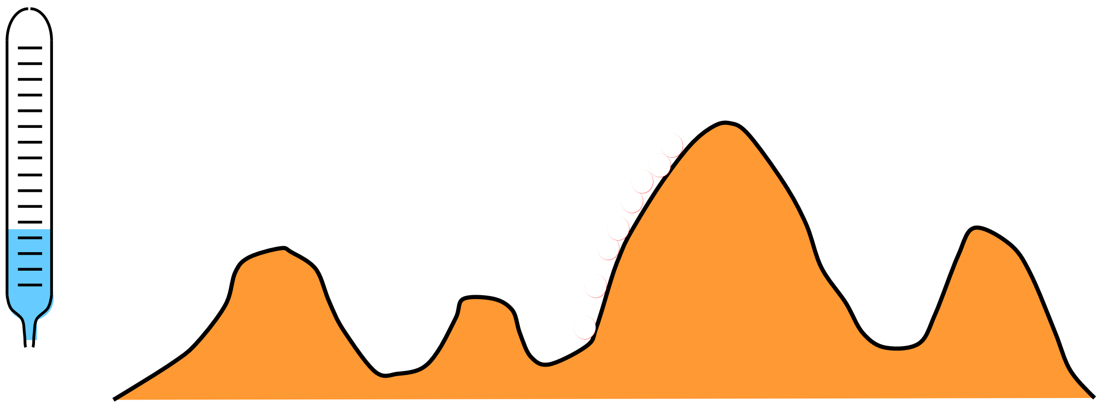
Simulierte Ausglühen



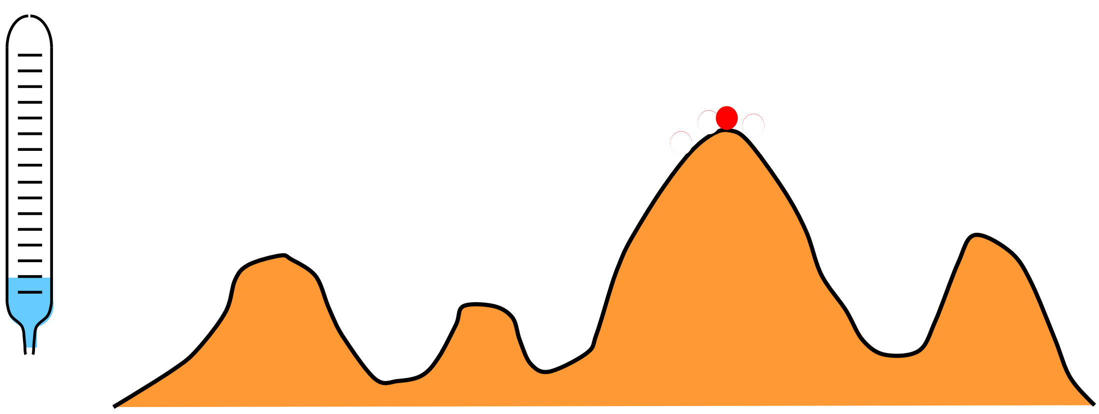
Simuliertes Ausglühen



Simuliertes Ausglühen



Simuliertes Ausglühen



Simuliertes Ausglühen

```
function Simuliertes-Ausglühen(Problem, Zeitplan) returns einen Lösungszustand
    inputs:           Problem   ; eine Problembeschreibung
                      Zeitplan  ; eine Abbildung der Zeit auf „Temperatur“
    local variables:  node      ; der aktuelle Knoten
                      next      ; ein Knoten
                      T         ; eine „Temperatur“, die die Wahrscheinlichkeit
                                          von Abwärtsschritten steuert
    node ← Make-Node(Initial-State[Problem])
    for t ← 1 to ∞ do
        T ← Zeitplan[t]
        if T = 0 then return node
        next ← ein zufällig gewählter Nachfolger von node
        ΔE ← Value[next] – Value[node]
        if ΔE > 0 then node ← next
              else node ← next nur mit Wahrscheinlichkeit   e^(ΔE/T)
end
```

t	1	2	3	…	20	21
T	20	19	18	…	1	0

Ist der neue Knoten besser, wird er auf jeden Fall genommen

Ist der neue Knoten schlechter, wird er mit einer gewissen Wahrscheinlichkeit trotzdem genommen

Vergleich: Suchen und Optimieren

	Problemlösen durch Suchen	Problemlösen durch Optimieren
Suchraum	Suche im **Zustandsraum**	Suche im **Lösungsraum**
Vorgehens-weise	Suche nach einer Lösung von einem Anfangszustand aus	Suche nach einer besseren Lösung von einer beliebigen Lösung aus
Operationen	Vordefinierte Aktionen	Veränderung einzelner Parameterwerte einer Lösung
Bewertung von Lösungen	Pfadkosten	Evaluierungsfunktion (Zielfunktion)
Ziel	Optimale Lösung	Optimale Lösung

Lokale Suche in kontinuierlichen Räumen

- Für einige Problemstellungen lässt sich eine Zielfunktion *f* definieren, die minimiert oder maximiert wird.
- Die optimalen Werte der Parameter der Zielfunktion sind gefunden, wenn die Funktion an dieser Stelle ein globales bzw. lokales Maximum (Minimum) besitzt.
- Oft werden auch Gleichungsnebenbedingungen gefordert.
- Numerische Verfahren nähern sich iterativ dem Ziel
- Bei ungünstigen Zielfunktionen (sehr flach und verrauscht) hilft nur eine zufällige Suche (Monte-Carlo-Methode).
- Oft wird der Gradient einer Funktion verwendet.

Lokale Suche in kontinuierlichen Räumen

$f : \Gamma \rightarrow R, \Gamma \subseteq R^N$ Zielfunktion, Kosten- oder Gütefunktion

$\min\limits_{x \in X} f(x)$ $x = (x_1, \ldots, x_N)^T \in R^N$

$X = \{x : c(x) = 0, h(x) \leq 0\}$ Optional: Nebenbedingungen

Gleichungsnebenbedingungen (GNB):

$c(x) = 0,\ c: R^N \rightarrow R^M$

(i.a. gilt $M<N$)

Ungleichungsnebenbedingungen (UNB):

$h(x) \leq 0,\ \ h: R^N \rightarrow R^Q$

Gleichungsnebenbedingungen

$c(x) = 0, \; c: R^N \to R^M \qquad x = (x_1, \dots, x_N)^T \in R^N$

$c(x) = (c_1(x), \dots, c_M(x))^T \qquad c_i : \Gamma \to R, \Gamma \subseteq R^N$

$$\boxed{c_i(x) = 0 \,, i = 1, \dots, M}$$

Ungleichungsnebenbedingungen

$h(x) \leq 0, \; h: R^N \to R^Q \qquad x = (x_1, \dots, x_N)^T \in R^N$

$h(x) = (h_1(x), \dots, h_Q(x))^T \qquad h_i : \Gamma \to R, \Gamma \subseteq R^N$

$$\boxed{h_i(x) \leq 0 \,, i = 1, \dots, Q}$$

Lokale und globale Minima

globales Minimum: $x^* \in X$

$$\boxed{f(x^*) \leq f(x), \qquad \forall x \in X}$$

lokales Minimum: $x^* \in X$

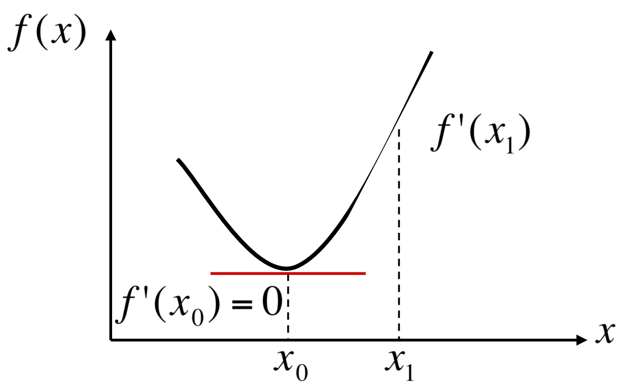
existiert Umgebung $U(x^*, \varepsilon) = \{z: \|z - x^*\| < \varepsilon\}, \qquad \varepsilon > 0$

$$\boxed{f(x^*) \leq f(x), \qquad \forall x \in U(x^*, \varepsilon) \cap X}$$

Lokale und globale Minima

Wir erinnern uns:

Um eine Funktion f zu minimieren, verwendet man die Ableitungsinformation. Der Vektor der Ableitungen von f in alle n Koordinatenrichtungen wird der Gradient von f genannt und oft als f ' geschrieben. In Analogie zum eindimensionalen Fall gilt auch für höherdimensionale Funktionen, dass bei einem Minimum die Ableitung gleich 0 sein muss, also die Bedingung f ' = 0.



Lokale und globale Minima

$$f(x) = f(x_1, \dots, x_N)$$

Gradient:

$$\frac{\partial f}{\partial x} = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \dots \\ \frac{\partial f}{\partial x_N} \end{pmatrix} = f_x = \nabla f = \nabla_x f$$

Minimierung ohne Nebenbedingungen

Problem: In der Regel ist die Funktion nicht lösbar oder liegt nicht einmal analytisch vor.

Man verwendet daher ein iteratives Verfahren, sucht also nach dem Minimum (Maximum).

<u>Algorithmus:</u>

1. Wahl des Startpunktes $\boldsymbol{x}^{(0)}$; Initialisierung des Iterationsindex auf l=0
2. Bestimmung der Suchrichtung $\boldsymbol{s}^{(l)}$
3. Bestimmung der skalaren Schrittweite $\eta^{(l)}>0$ und Neuberechnung von $\boldsymbol{x}$ mit
$$x^{(l+1)} = x^{(l)} + \eta^{(l)} \cdot s^{(l)}$$
4. Bei Erfüllung des Abbruchkriteriums: Ende; Adernfalls zu (2) mit $l=l+1$

Gradientenverfahren (steilster Abstieg)

Für den Gradienten gilt die geometrisch interessante Eigenschaft, dass der negative Gradient in Richtung des steilsten Abstiegs zeigt und somit in einer hinreichend kleinen Umgebung des Punktes $\boldsymbol{x}_0$ gilt:

$$f(x_0 - \eta \cdot f'(x_0)) \leq f(x_0 + \mathrm{v})$$

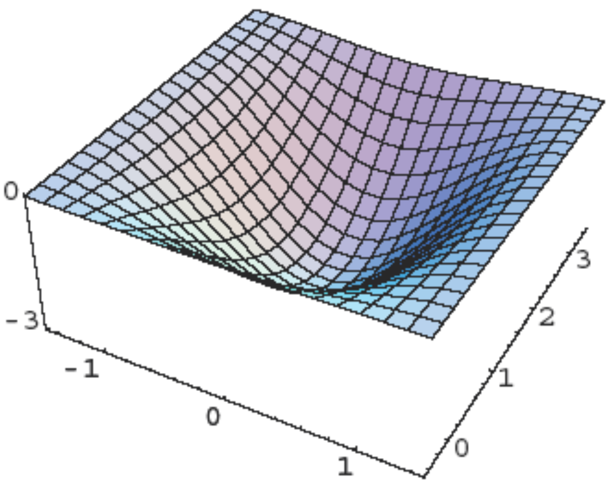
wobei ν ein beliebiger anderer Vektor derselben Länge ist.

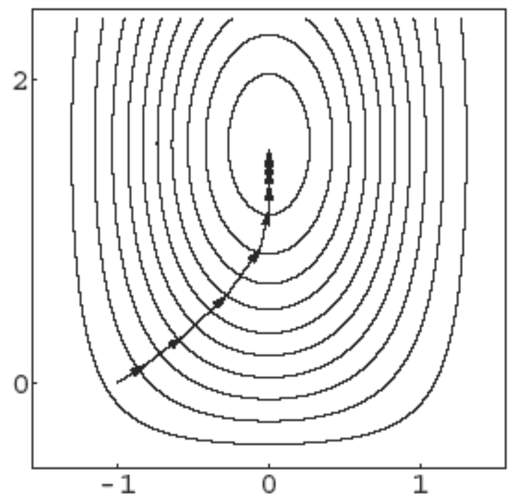
Gradientenverfahren

Das Prinzip des Gradientenabstiegs zur Minimumsfindung nutzt diese Eigenschaft aus, um iterativ von einem Startwert nach unten in Richtung des Minimums zu springen. Jeweils neue Zwischenresultate $\boldsymbol{x}^{neu}$ werden aus den alten $\boldsymbol{x}^{alt}$ durch die Formel

$$x^{neu} = x^{alt} - \eta \cdot f'\left(x^{alt}\right) \qquad \text{oder} \qquad \Delta x = x^{neu} - x^{alt} = -\eta \cdot f'\left(x^{alt}\right)$$

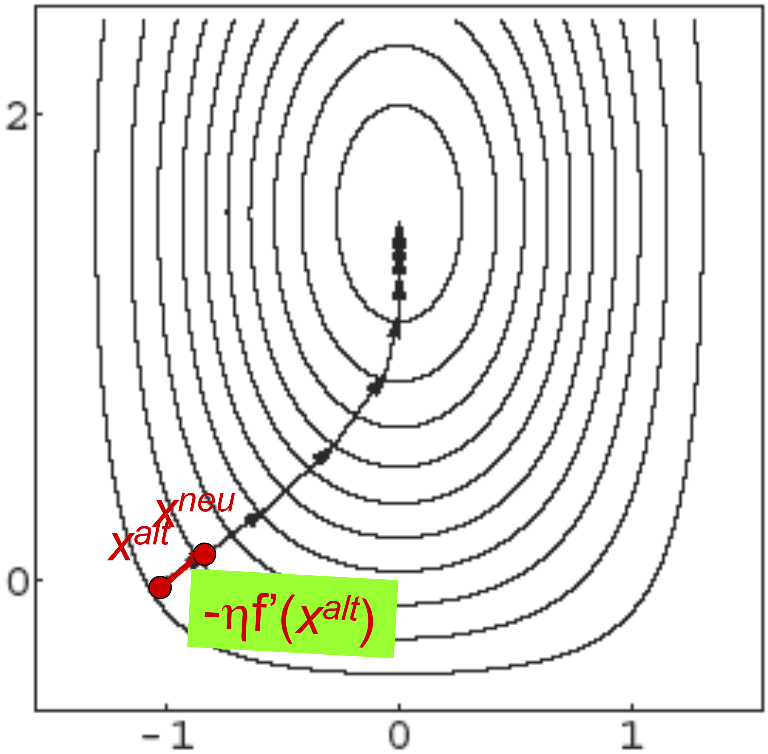
berechnet. η wird als Schrittweite oder Lernrate bezeichnet.





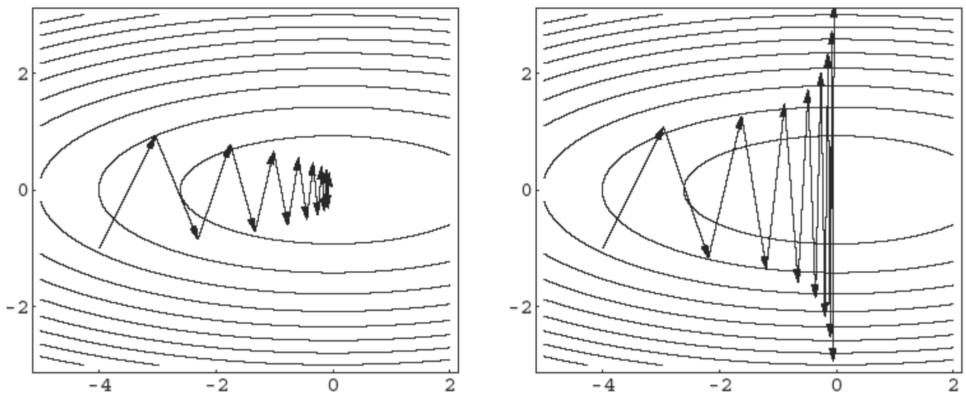
Gradientenverfahren

$$x^{neu} = x^{alt} - \eta \cdot f'(x^{alt})$$



Gradientenverfahren

- Bei Wahl einer zu großen Lernrate η kann es allerdings zu starken Fluktuationen oder gar zu keiner Konvergenz des Lernes kommen.



- Bei Wahl einer zu kleinen Lernrate η kann das Verfahren in einem lokalen Minimum hängenbleiben.
- Bei kleinen Gradienten kann eine langsame Konvergenz die Folge sein.

Minimierung unter Gleichungsnebenbedingungen

$$\min_{x \in X} f(x) \qquad x = (x_1, \dots, x_N)^T \in R^N$$

$$X = \{x : c(x) = 0\}$$ zulässiger Bereich

$$c_i(x) = 0\,, i = 1, \dots, M$$

Lagrangefunktion:

$$\boxed{L(x,\lambda) = f(x) + \lambda^T c(x) = f(x) + \sum_{i=1}^{M} \lambda_i \cdot c_i(x)}$$

$\lambda = (\lambda_1, \dots, \lambda_M)^T$ Lagrangemultiplikatoren

Diese Methode führt eine neue unbekannte skalare Variable λ für jede Nebenbedingung ein, und definiert eine Linearkombination, die die Multiplikatoren als Koeffizienten einbindet. Das reduziert das Nebenbedingungsproblem auf ein Problem ohne Nebenbedingung.

Notwendige Optimalitätsbedingungen (lokales Minimum)

Lagrangefunktion:

$$L(x,\lambda) = f(x) + \lambda^T c(x) = f(x) + \sum_{i=1}^{M} \lambda_i \cdot c_i(x)$$

$$L_x(x^*,\lambda^*) = f_x(x^*) + \mathbf{c}_x^T(x^*) \cdot \lambda^* = 0$$

Optimalitätsbedingungen:

$$\boxed{f_x(x^*) + \sum_{i=1}^{M} \lambda_i^* \cdot c_{i,x}^T(x^*) = 0}$$

$$\boxed{L_\lambda(x^*,\lambda^*) = \mathbf{c}(x^*) = 0}$$

Jacobi – Matrix (Ableitungsmatrix)

$$\frac{\partial c}{\partial x} = c_x = \begin{pmatrix} \frac{\partial c_1}{\partial x_1} & \cdots & \frac{\partial c_1}{\partial x_N} \\ \vdots & & \vdots \\ \frac{\partial c_M}{\partial x_1} & \cdots & \frac{\partial c_M}{\partial x_N} \end{pmatrix} \qquad \frac{\partial \boldsymbol{c}}{\partial x} = J\boldsymbol{c}$$

$$\frac{\partial \boldsymbol{c}}{\partial x} = \begin{pmatrix} - & c_{1,x} & - \\ \vdots & & \vdots \\ - & c_{M,x} & - \end{pmatrix} \qquad c_{i,x} = \left(\frac{\partial c_i}{\partial x_1}, \ldots, \frac{\partial c_i}{\partial x_N}\right)$$

Beispiel

$$f : \Gamma \to R, \Gamma \subseteq R^2 \qquad \Gamma = \{(x_1, x_2): x_1 \geq 0, x_2 \geq 0\}$$

$$\boxed{\min_{x \in X} f(x) = x_1^2 + x_2^2} \qquad x = (x_1, x_2)^T$$

Nebenbedingung:

$$c_1(x) = x_1 + x_2 - 10$$

Ansatz:

$$L(x_1, x_2, \lambda) = x_1^2 + x_2^2 + \lambda(x_1 + x_2 - 10)$$

Beispiel

$$L(x_1, x_2, \lambda) = x_1^2 + x_2^2 + \lambda(x_1 + x_2 - 10)$$

$$\left.\boxed{\begin{aligned} L_{x_1}(x_1, x_2, \lambda) &= 2x_1 + \lambda = 0 \\ L_{x_2}(x_1, x_2, \lambda) &= 2x_2 + \lambda = 0 \\ L_{\lambda}(x_1, x_2, \lambda) &= x_1 + x_2 - 10 = 0 \end{aligned}}\right\} \quad \begin{aligned} x_1 &= x_2 \\ \lambda &= -10 \\ x_1 &= x_2 = 5 \end{aligned}$$

Minimierung unter Gleichungs- und Ungleichungsnebenbedingungen

$$\boxed{\min_{x \in X} f(x)} \qquad x = (x_1, \ldots, x_N)^T \in R^N$$

$$\boxed{X = \{x : c(x) = 0, h(x) \leq 0\}}$$ zulässiger Bereich

$$c_i(x) = 0 \, , i = 1, \ldots, M$$

$$h_i(x) \leq 0 \, , i = 1, \ldots, Q$$

Allgemeine Lagrangefunktion

$$L(x, \lambda, \mu) = f(x) + \lambda^T \boldsymbol{c}(x) + \mu^T \boldsymbol{h}(x)$$

$\lambda = (\lambda_1, \dots, \lambda_M)^T$ Lagrangemultiplikatoren

$\mu = (\mu_1, \dots, \mu_Q)^T$ Kuhn-Tucker-Multiplikatoren

Notwendige Optimalitätsbedingungen (lokales Minimum)

Lagrangefunktion :

$$L(x, \lambda, \mu) = f(x) + \lambda^T \boldsymbol{c}(x) + \mu^T \boldsymbol{h}(x)$$

Optimalitätsbedingungen:

$$\boxed{L_x(x^*, \lambda^*, \mu^*) = f_x(x^*) + \boldsymbol{c}_x^T(x^*) \cdot \lambda^* + \boldsymbol{h}_x^T(x^*) \cdot \mu^* = 0}$$

$$\boxed{L_\lambda(x^*, \lambda^*, \mu^*) = \boldsymbol{c}(x^*) = 0}$$

$$\mu^* L_\mu(x^*, \lambda^*, \mu^*) = \mathbf{0} \Longleftrightarrow \boxed{\boldsymbol{h}(x^*)^T \cdot \mu^* = 0}$$

$$\boxed{\mu^* \geq 0}$$

$$\boxed{\boldsymbol{h}(x^*) \leq 0}$$

Add

Bemerkung

$$h(x^*)^T \cdot \mu^* = 0$$ kann geschrieben werden als

$$\boxed{h_i(x^*)^T \cdot {\mu_i}^* = 0 \ \ , i = 1, \dots, Q}$$

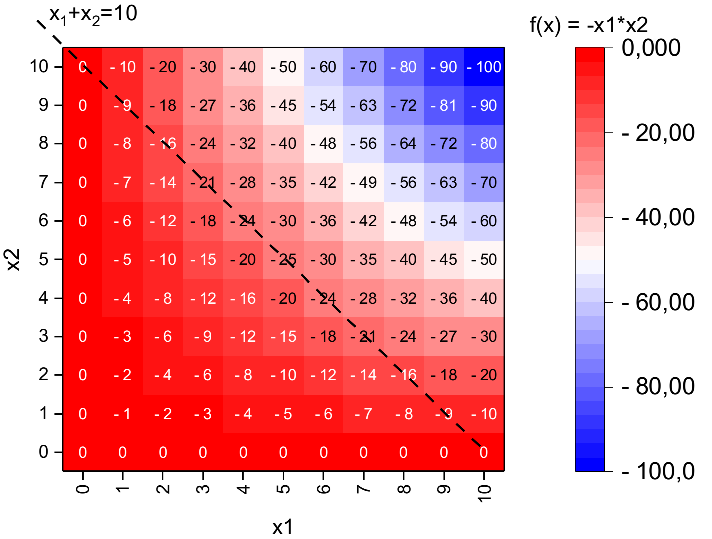
wegen $$h_i(x^*) \leq 0, \mu_i \geq 0$$

Beispiel

$$f: \Gamma \rightarrow R, \Gamma \subseteq R^2 \qquad \Gamma = R^2$$

$$\boxed{\min_{x \in X} f(x) = -x_1 x_2}$$

$$x = (x_1, x_2)^T$$



Nebenbedingungen:

$$c_1(x) = x_1 + x_2 - 10$$

$$h_1(x) = -x_1$$

$$h_2(x) = -x_2$$

Add

Beispiel

Ansatz:

$$L(x_1, x_2, \lambda, \mu_1, \mu_2) = x_1 x_2 + \lambda(x_1 + x_2 - 10) - \mu_1 x_1 - \mu_2 x_2$$

$$L_{x_1}(x_1, x_2, \lambda, \mu_1, \mu_2) = -x_2 + \lambda - \mu_1 = 0$$

$$L_{x_2}(x_1, x_2, \lambda, \mu_1, \mu_2) = -x_1 + \lambda - \mu_2 = 0$$

$$L_{\lambda}(x_1, x_2, \lambda, \mu_1, \mu_2) = x_1 + x_2 - 10 = 0$$

$$\boxed{\boldsymbol{h}(x^*) \leq 0} \qquad \boxed{\mu^* \geq 0}$$

$$-x_1 \leq 0 \qquad -x_2 \leq 0 \qquad \mu_1, \mu_2 \geq 0$$

$$-x_1 \mu_1 = 0 \qquad -x_2 \mu_2 = 0 \qquad \boxed{\boldsymbol{h}(x^*)^T \cdot \mu^* = 0}$$

Beispiel

Im Lösungsraum müssen folgende Fälle betrachtet werden:

1. Lösung

$$\mu_1, \mu_2 = 0$$

3. Lösung

$$\mu_1 = 0; \mu_2 > 0$$

2. Lösung

$$\mu_1 > 0; \mu_2 > 0$$

4. Lösung

$$\mu_1 > 0; \mu_2 = 0$$

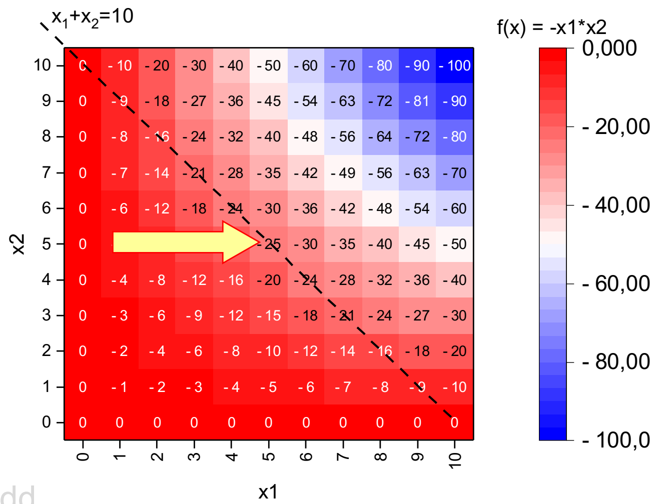
Add

Beispiel

1. Lösung

$$-x_2 + \lambda - \mu_1 = 0$$
$$-x_1 + \lambda - \mu_2 = 0$$
$$x_1 + x_2 - 10 = 0$$

$$\mu_1, \mu_2 = 0 \quad \Rightarrow \quad x_1 = x_2 = 5$$



Add

lokales Minimum

$$f(x) = -x_1 x_2 = -25$$

Beispiel

1. Lösung

$$\mu_1, \mu_2 = 0 \quad \Rightarrow \quad x_1 = x_2 = 5$$

lokales Minimum

2. Lösung

$$\mu_1 \neq 0 \vee \mu_2 \neq 0 \quad \Rightarrow \quad x_1 = 0 \vee x_2 = 0$$

$$\Rightarrow \quad f(x) = -x_1 x_2 = 0$$

globales Maximum

Beispiel

3. Lösung $\mu_1 = 0; \mu_2 > 0$

$$-x_2 + \lambda - \mu_1 = 0 \quad (1)$$
$$-x_1 + \lambda - \mu_2 = 0 \quad (2)$$
$$x_1 + x_2 - 10 = 0 \quad (3)$$

$(1) mit\ \mu_1 = 0$ $$-x_2 + \lambda = 0 \quad (4)$$

$(4) in (2)$ $$-x_1 + x_2 - \mu_2 = 0 \quad (5)$$
$$x_1 + x_2 - 10 = 0 \quad (3)$$

$(5) + (3)$ $$2x_2 - \mu_2 - 10 = 0 \quad (6)$$

$Umstellen\ von\ (6)$ $$x_2 = \frac{10 + \mu_2}{2} \quad (6)'$$

$Umstellen\ von\ (3)$ $$x_1 = 10 - x_2 \quad (3)'$$

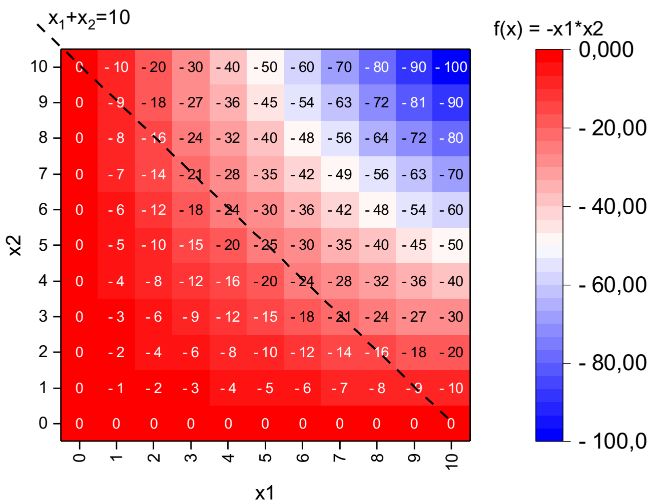
$(6)' in (3)'$ $$x_1 = 10 - \frac{10 + \mu_2}{2} \quad (7)$$

$$x_1 = 5 - \frac{\mu_2}{2} \quad (7)$$

Add

Beispiel

3. Lösung $\mu_1 = 0; \mu_2 > 0$



$$x_1 = 5 - \frac{\mu_2}{2} \quad (7)$$

$$\mu_2 > 0 \rightarrow x_1 < 5$$

$$x_1 \geq 0 \rightarrow \mu_2 \leq 10$$

$$0 < \mu_2 \leq 10$$

$$0 \leq x_1 < 5$$

$$f(x) = -x_1 x_2 = 0$$

$$10 \cdot 0 \leq f(x) < -5 \cdot 5 = -25$$

globales Maximum

$x_1 = 0, x_2 = 0$

lokales Minimum

$x_1 < 5, x_2 > 5$

Add

Beispiel

Im Lösungsraum müssen folgende Fälle betrachtet werden:

1. Lösung

$$\mu_1, \mu_2 = 0$$

lokales Minimum

$$f(x) = -x_1 x_2 = -25$$

2. Lösung

$$\mu_1 = 0; \mu_2 > 0$$

globales Maximum

$$f(x) = -x_1 x_2 = 0$$

3. Lösung

$$\mu_1 > 0; \mu_2 = 0$$

globales Maximum

$$f(x) = -x_1 x_2 = 0$$

4. Lösung

$$\mu_1 > 0; \mu_2 > 0$$

globales Maximum

$$f(x) = -x_1 x_2 = 0$$

Add

Beispiel

4. Lösung $\mu_1 > 0; \mu_2 > 0$

$$-x_2 + \lambda - \mu_1 = 0 \ (1)$$
$$-x_1 + \lambda - \mu_2 = 0 \ (2)$$
$$x_1 + x_2 - 10 = 0 \ (3)$$

$(1) - (2)$ $$-x_2 + x_1 - \mu_1 + \mu_2 = 0 \ (4)$$

$(4) - (3)$ $$x_1 + x_2 - 10 = 0 \ (3)$$

$$-2x_2 - \mu_1 + \mu_2 + 10 = 0 \ (5)$$

$(5) + (3)$

$Umstellen\ von\ (5)$ $$x_2 = \frac{-\mu_1 + \mu_2 + 10}{2} \ (5)'$$

$Umstellen\ von\ (3)$ $$x_1 = 10 - x_2 \ (3)'$$

$(5)'in(3)'$ $$x_1 = 10 - \frac{-\mu_1 + \mu_2 + 10}{2} (6)$$

$$x_1 = 5 - \frac{-\mu_1 + \mu_2}{2} (6)$$

Beispiel

4. Lösung $\mu_1 > 0; \mu_2 > 0$

$$x_1 = 5 - \frac{-\mu_1 + \mu_2}{2} \quad (6)$$

$$-x_1 \leq 0 \qquad x_1 \leq 0 \qquad 5 - \frac{-\mu_1 + \mu_2}{2} < 0$$

$$Umstellen \qquad 10 < -\mu_1 + \mu_2$$

$$Umstellen \qquad 10 + \mu_1 < \mu_2$$

$$\mu_2 > \mu_1 + 10 \quad (7)$$

$$(7) in\ (6) \quad x_1 = 5 - \frac{-\mu_1 + \mu_1 + 10}{2}$$

$$x_1 = 5 - 5 = 0\ (6)$$

Add $f(x) = -x_1 x_2 = 0$