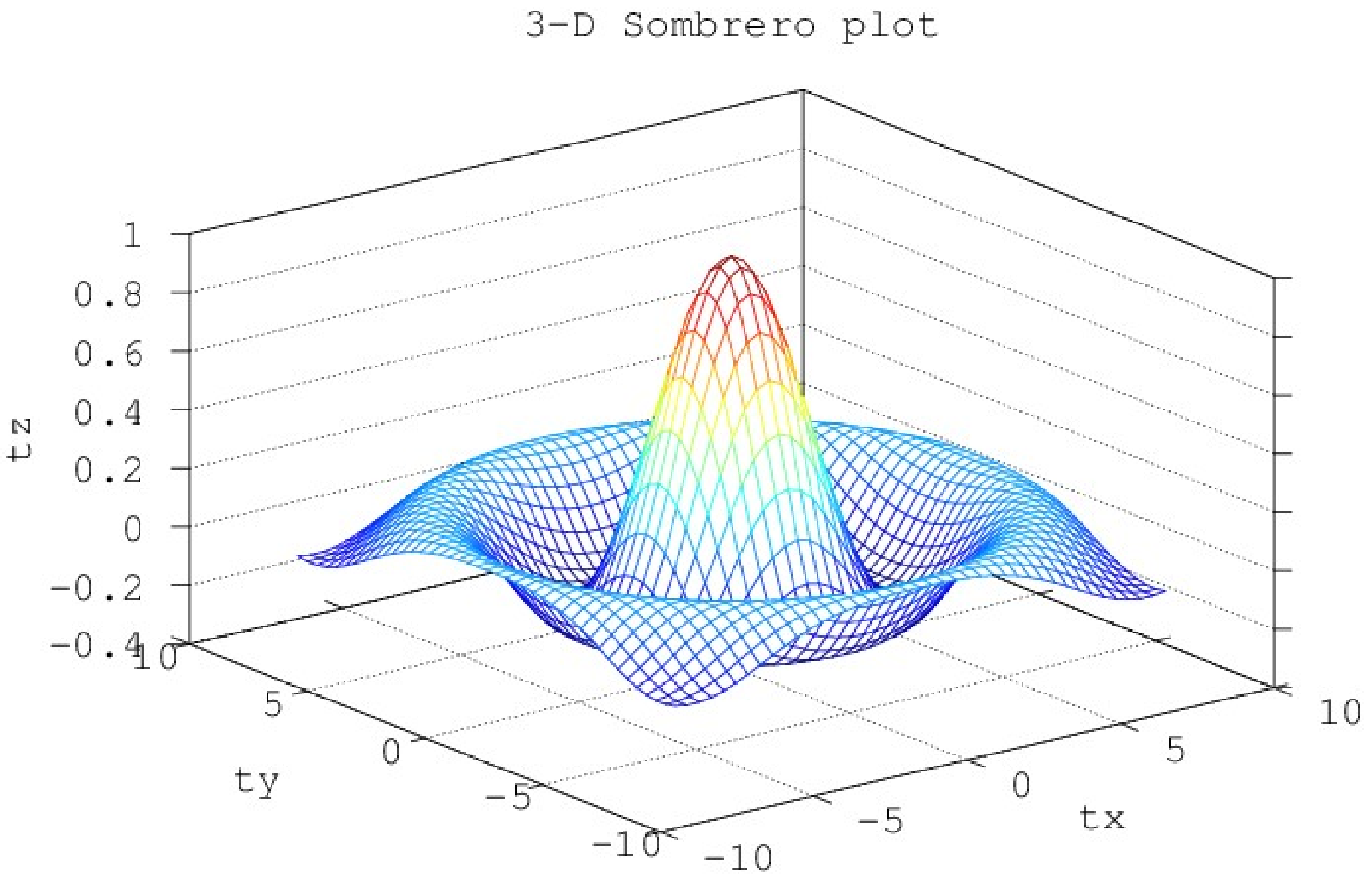
# Mehrdimensionale Optimierung ohne Nebenbedingungen

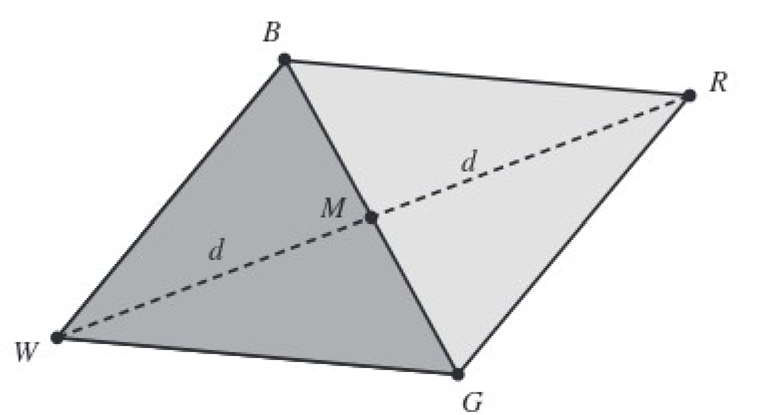


# Nelder-Mead-Verfahren

- Geometrischer Ansatz: ein Simplex (im zweidimensionalen Fall ein Dreieck) „taumelt“ die Funktion herab und zieht sich um das Minimum zusammen.
- Verwaltet drei Punkte: B (*best*), G (*good*) und W (*worst*). Diese bilden den Simplex BGW.
- Iteration besteht aus 3 Phasen: Ordering, Centroid und Transformation.
- Ordering = Sortiere Punkte.
- Centroid = Berechne Mittelpunkt.
- Transformation = Transformiere Simplex.

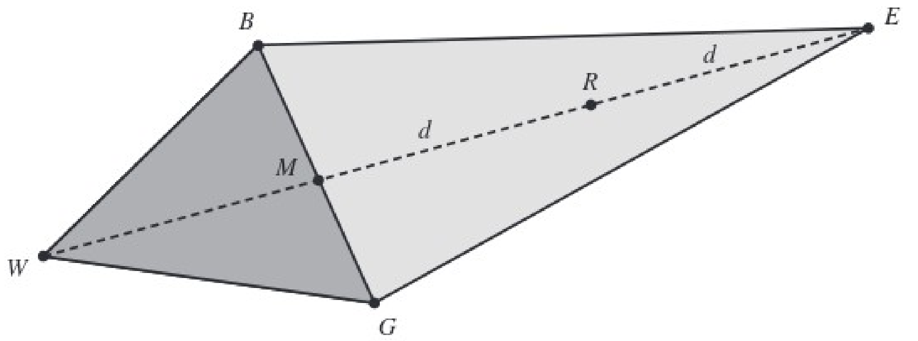
# Ablauf: Nelder-Mead-Verfahren

1) Wähle 3 Startpunkte – diese bilden den Simplex.
2) Sortiere diese anhand ihrer Funktionswerte
3) Berechne den Mittelpunkt der zwei besten Punkte m
4) Reflektiere den schlechtesten Punkt am Mittelpunkt m. Nenne diesen Punkt r.



# Ablauf: Nelder-Mead

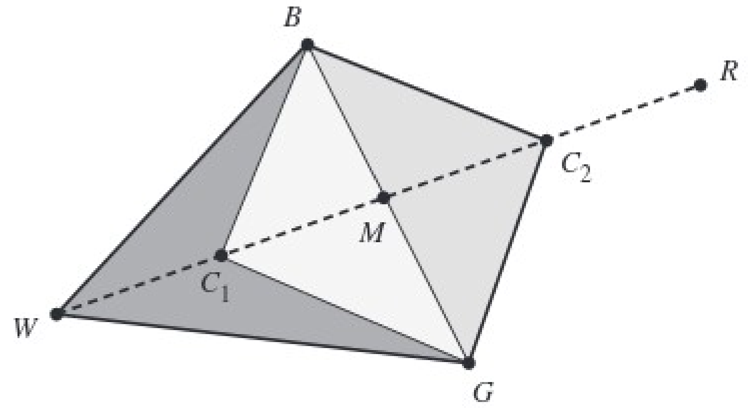
5) Wenn r besser als der beste Punkt bis jetzt ist, haben wir etwas richtig gemacht. Berechne den expandierten Punkt e anhand des schlechtesten Wertes.




6) Ersetze den schlechtesten Wert („akzeptiere“) durch den besseren der Punkte (e, r) und gehe zu (2).

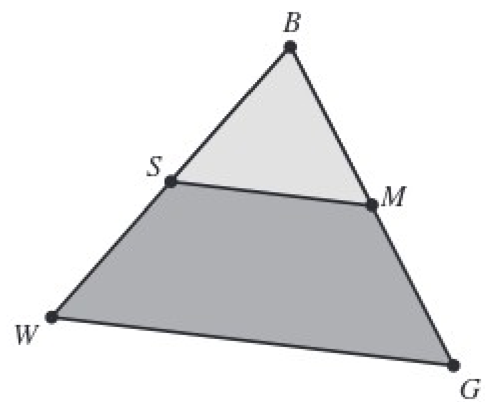
# Ablauf: Nelder-Mead

7) Wenn r besser als der schlechteste Punkt ist, berechne den kontrahierten Punkt c1. Ist dieser besser als r, akzeptiere c1. Gehe zu (2). Wenn nicht, berechne den kontrahierten Punkt c2. Ist dieser besser als W, akzeptiere c2. Gehe zu (2).



# Ablauf: Nelder-Mead

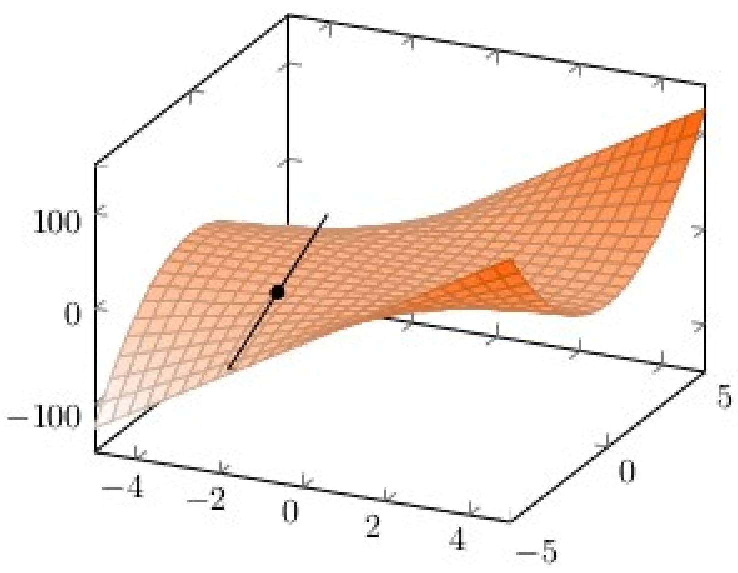
8) Komprimiere die beiden schlechtesten Punkte G und W in Richtung B. Gehe zu (2).




Kumulative Quelle:
http://mathfaculty.fullerton.edu/mathews/n2003/neldermead/NelderMeadProof.pdf

# Gradientenverfahren

- Gradient $\nabla f(x,y)$ zeigt immer in die Richtung des steilsten Anstiegs (Negiert: in die Richtung des steilsten Abfalls)
- Idee: verwende diese Eigenschaft zur Optimierung: wandere die Funktion entlang in diese Richtung (Gradient wird dabei kürzer), bis keine Änderung mehr feststellbar ist



# Ablauf: Gradientenverfahren

1) Wähle beliebigen Startpunkt $P=(x,y)$
2) Bilde die erste Ableitung, also den Gradienten
3) Ist $\nabla f(x,y)=0$ , haben wir eine Extremstelle gefunden.
4) Setze P in den Gradienten ein und negiere
5) Bestimme, wie weit in diese Richtung gewandert werden soll → *line-search*

# Line search

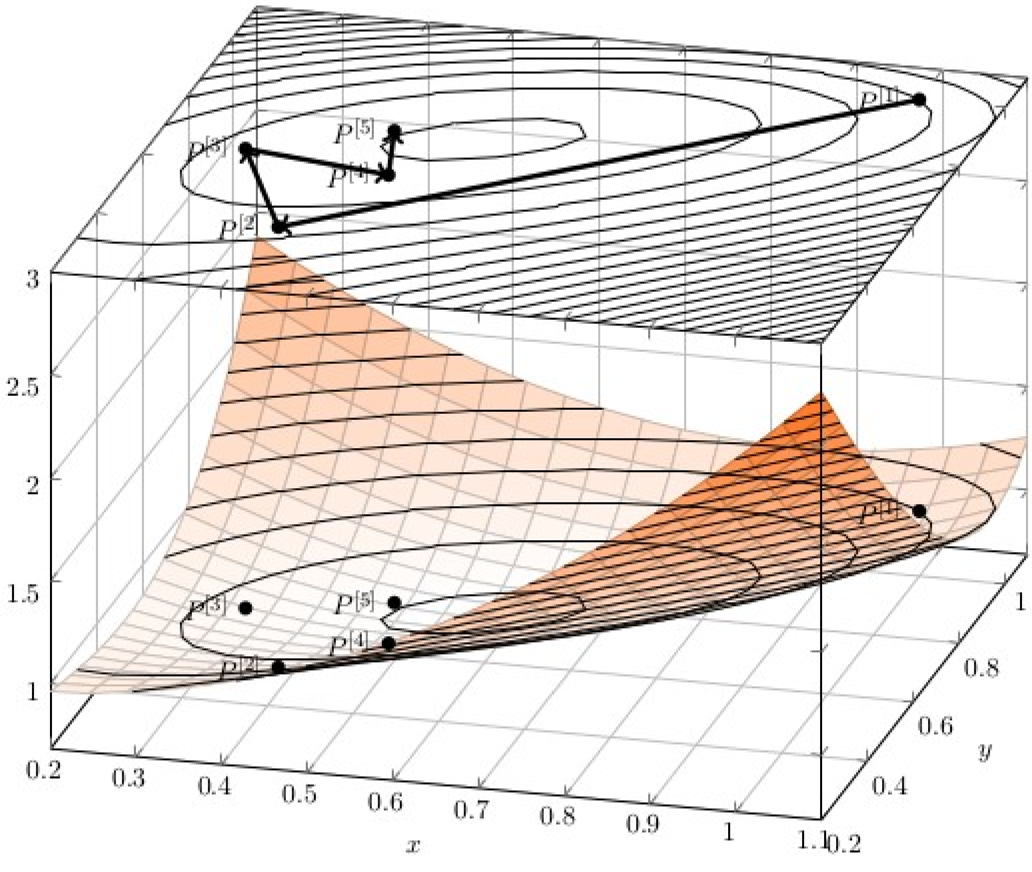
- Exakte Line-Search unnötig aufwändig, stattdessen → Heuristik
- Armijo-Bedingung: $\varphi(\alpha) \leqslant \varphi(0) + c \cdot \alpha^{[n]} \cdot \varphi(0)'$

1) Suche ein α, für welches diese Ungleichung erfüllt ist

2) Wenn erfüllt, wurde eine passende Schrittweite gefunden und ein neuer Punkt kann berechnet werden. Fahre fort bei (3) auf der vorigen Folie.

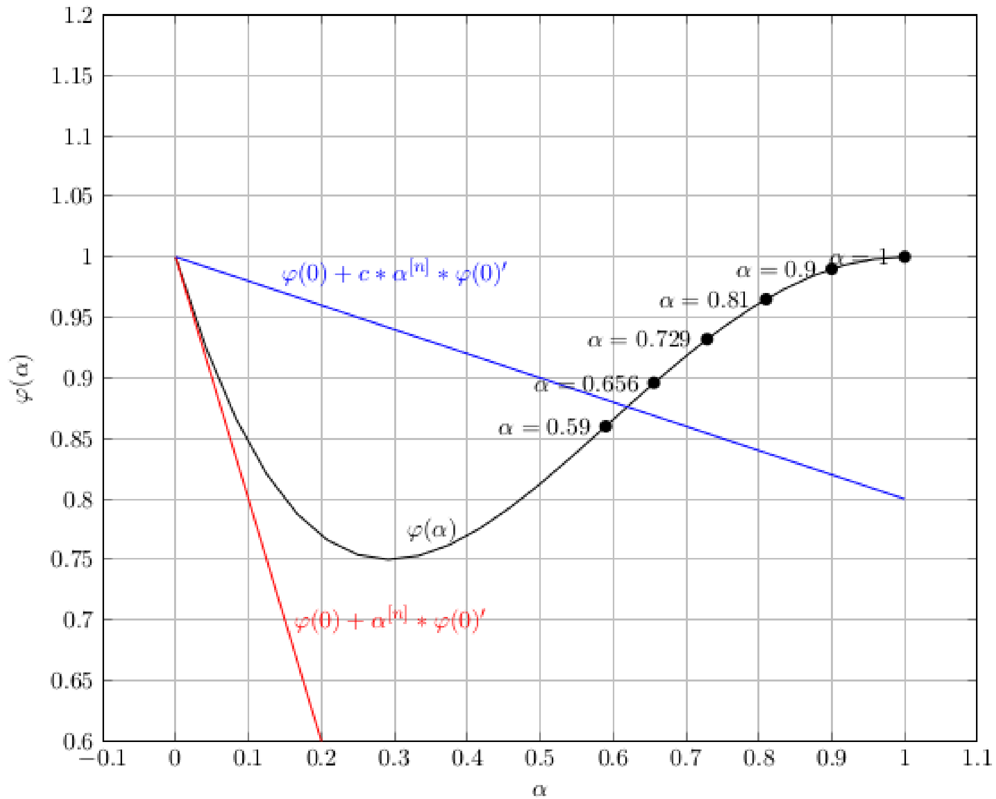
# Beispiel

- Zu minimierende Funktion: $f(x,y)=y^4+2x^2-3xy+1$
- Abbruchbedingung: $\|\nabla f(x,y)\|\leq 0.3$



# Beispiel

- Startpunkt: P = (1, 1)
- Ableitung: $\nabla f(x,y) = \begin{pmatrix} 4x - 3y \\ 4y^3 - 3x \end{pmatrix} = \vec{g}$
- Steepest descent in Richtung: $-\nabla f(1,1) = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$
- Line search



# Beispiel

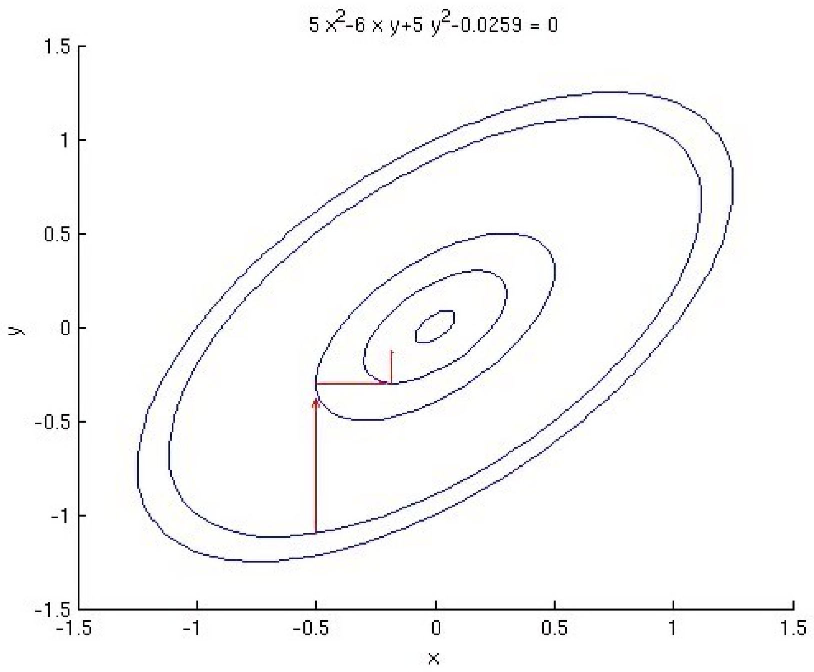
- Gültiges α: 0.59 (nach 6 Versuchen)
- Wandere um diese Schrittlänge weiter:

$$P^{[2]} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 0.59 \cdot \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0.41 \\ 0.41 \end{pmatrix}$$

- Bilde wieder den Gradienten, wiederhole
- Abbruch sobald Abbruchbedingung erfüllt

# Koordinatenabstiegsmethode

- Minimiere immer nach einer Achse
- Neuer Punkt wird beibehalten, eingesetzt – nur die Koordinate, nach der optimiert werden soll, bleibt variabel



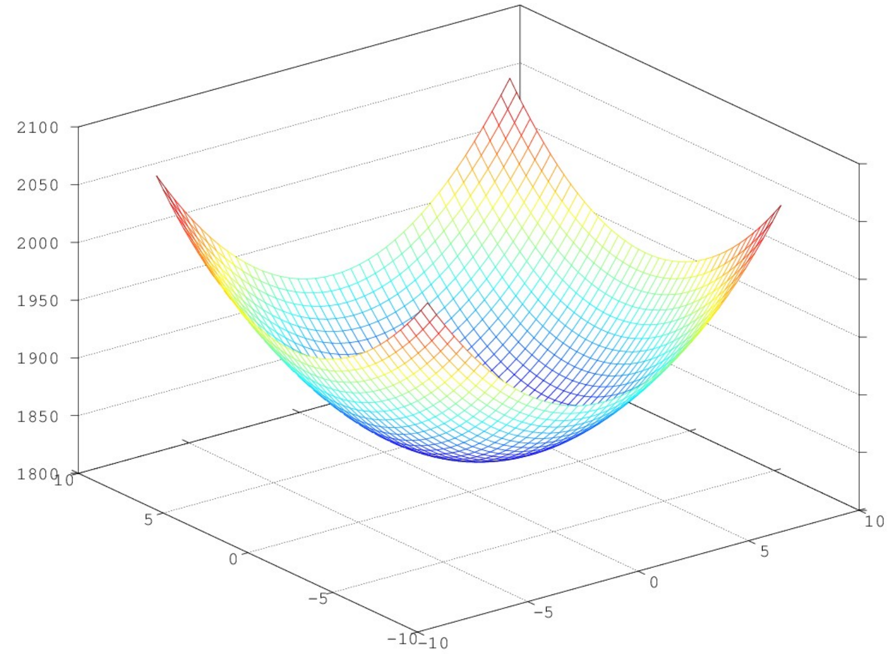

Quelle: de.wikipedia

# Ablauf: Koordinatenabstiegsmethode

1) Wähle zufälligen Startpunkt
2) Wähle erste Koordinate, nach der minimiert werden soll
3) Setze für andere Koordinate Wert aus derzeitigem Punkt ein
4) Minimiere eindimensionale Funktion
5) Setze gefundenes Minimum in Punkt ein
6) Wähle andere Koordinate, gehe zu (3)

# Beispiel

- Funktion $f(x,y)=2x^2+2xy+1.5y^2$
- Wähle Startpunkt (1,1)
- Setze für x ein: $x'=1+d$
- Setze für y ein: $y'=1+0d$
- Minimiere nach d,
- Minimum bei $\left(\frac{-3}{2}\right)$, neuer Punkt: $P=\begin{pmatrix}1-\frac{3}{2}\\1\end{pmatrix}=\begin{pmatrix}\frac{-1}{2}\\1\end{pmatrix}$
- Wiederhole mit y, etc.



# To Do

- Theorieausarbeitungen zusammentragen, überarbeiten und ergänzen
- 2 / 3 Implementierungen (Coordinate-Descent und Gradientenverfahren) ausstehend
- Visualisierungsmethoden erschließen
- Homepage: Weitere interessante Beispiele, Text, Design