

Thema: Abstiegsverfahren

Gegeben: Quadratisches Funktional.

$$Q(\vec{x}) = x^T A x + 2b^T x + c$$

$A \in \mathbb{R}^{n \times n}$ ist symmetrisch positiv definit.

$b \in \mathbb{R}^n$

gesucht.: minimalstelle

Methode gradient descent:

$$x_{i+1} = x_i - t_i \cdot \nabla f(x_i)$$

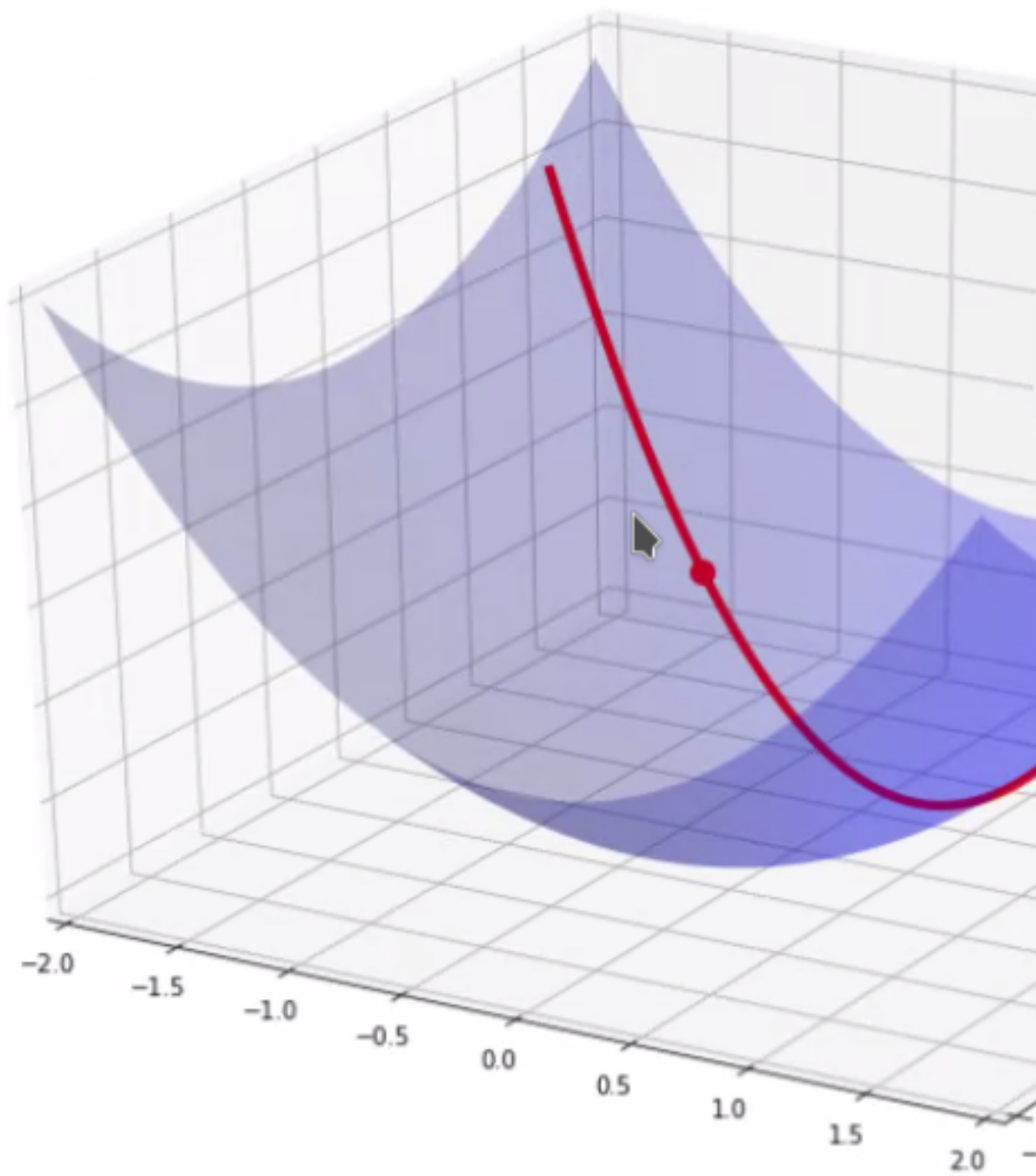
optimale schrittweite τ , sodass $x_1 = x_0 + \tau \cdot s_0$

suche τ , sodass $Q(x_1)$ minimal wird.

wir haben also ein eindimensionales optimierungsproblem in τ .

$$q(\tau) = Q(x_0 + \tau \cdot s_0)$$

wir wissen, dass $q(\tau)$ eine 1-d parabel entlang des gradienten auf der quadratischen Funktion ist



$$q(\tau)d/d\tau = d/d\tau Q(x_0 + \tau \cdot s_0)$$

Kettenregel

$$q(\tau)d/d\tau = dQ/dx \cdot Q(x_0 + \tau \cdot s_0) \cdot d/d\tau (x_0 + \tau \cdot s_0)$$

$$(\nabla Q(x_0 + \tau \cdot s_0))^T \cdot s_0$$

$$\langle \nabla Q(x_0 + \tau \cdot s_0), s_0 \rangle$$

der Gradient steht senkrecht auf der Suchrichtung: der Erste komponent ist präzise das, was wir optimieren wollen!

$$\nabla Q(x) = 2Ax + 2b$$

also folgt für τ

$$\langle 2(A(x_0 + \tau s_0) + b), s_0 \rangle$$

$$\langle 2(Ax_0 + A\tau s_0 + b), s_0 \rangle$$

$$2((Ax_0)^T s_0 + \tau(A s_0)^T s_0 + b^T s_0) = 0$$

$$\iff \tau(A s_0)^T s_0 = -((Ax_0)^T s_0 + b^T s_0)$$

$$\iff \tau = -\frac{\langle Ax_0 + b, s_0 \rangle}{\langle A s_0, s_0 \rangle}$$

Für Tau betrachten wir also das verhältniss vom inneren produkt der suchrichtung zur 1. ableitung des momentanen punktes und der Suchrichtung und 2. ableitung.

gradientenverfahren: Suchrichtung ist Gradient.

A konjugation:

bisher: gehen vom gradienten im schnellsten weg zur nächsten niveaulinie. $\langle s_1, s_0 \rangle = 0$ für den nächsten wert.

A konjugation ändert das krit. auf $\langle s_1, A s_0 \rangle = 0$

für uns hier $s_1^T (-4, -8)^T = 0 \implies s_1 = (2, -1)^T$

f)

1. Schritt gleich gradientenverfahren:

$$x_1 = x_0 + \tau_0 s_0 = (-1, 1)^T$$

2. Schritt

wir gehen in richtung des konj. Gradienten. (also hier das zuvor berechnete s_1)

$$\tau_1 = \frac{\langle Ax_1 + b, s_1 \rangle}{\langle A s_1, s_1 \rangle} = 1/7$$

$$y_2 = (-5/7, 6/7)^T$$

WIR SIND AM NULLPUNKT (weil wir exakt rechnen und quadratisch sind)

Reading zu CG:

“conjugate gradients without the agonizing pain”