

F9

- Optimering, NAM Kap. 7

$\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x})$ utan bivillkor

$\min_{\mathbf{L} < \mathbf{x} < \mathbf{U}} f(\mathbf{x})$ enkla gränser

$\min_{\mathbf{g}(\mathbf{x})=0} f(\mathbf{x})$ med likhetsbivillkor

$\min_{\mathbf{h}(\mathbf{x}) < 0} f(\mathbf{x})$ med olikhetsbivillkor

f : objektsfunktion

Optimering, II

Exempel

$$\min_{\mathbf{x} \in \mathbb{R}^2} \sum_j |\mathbf{x} - \mathbf{x}_j|$$

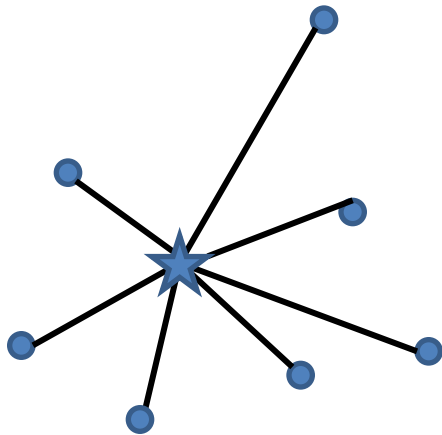
$$\min_{(x,y)} f(x,y) = \min_{(x,y)} \sum_j \sqrt{(x - x_j)^2 + (y - y_j)^2}$$

$$f(x+h, y+k) = f(x,y) + \frac{\partial f}{\partial x} h + \frac{\partial f}{\partial y} k + \dots$$

$$\frac{1}{2} \left(\frac{\partial^2 f}{\partial x^2} h^2 + 2 \frac{\partial^2 f}{\partial x \partial y} h k + \frac{\partial^2 f}{\partial y^2} k^2 \right) + \dots$$

$$f(x,y) + \mathbf{g}^T \mathbf{h} + \frac{1}{2} \mathbf{h}^T \mathbf{H} \mathbf{h} + \dots,$$

$$\mathbf{h} = \begin{pmatrix} h \\ k \end{pmatrix}, \mathbf{g} = \text{gradient} = \begin{pmatrix} f_x \\ f_y \end{pmatrix}, \mathbf{H} = \text{Hessian} = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix}.$$

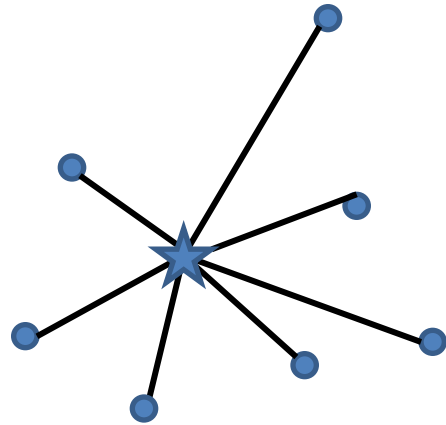


Optimering, III

$\mathbf{g} = 0$: nödvändigt för stationär punkt

$\mathbf{g} = 0, \mathbf{H}$ positivt definit ($\mathbf{h}^T \mathbf{H} \mathbf{h} > 0$ för alla $\mathbf{h}$) tillräckligt för lokalt minimum :

$f(x, y) < f(x + h, y + k)$ i omgivning $0 < \|\mathbf{h}\| < \delta$



$$\sum_j r_j, r_j = \sqrt{(x - x_j)^2 + (y - y_j)^2}$$

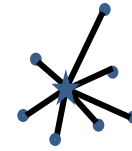
$$\mathbf{g} = \sum_j \mathbf{g}_j, \mathbf{g}_j = \frac{1}{r_j} (\mathbf{x} - \mathbf{x}_j)$$

$$\mathbf{H} = \sum_j \mathbf{H}_j, \mathbf{H}_j = \frac{1}{r_j^3} \begin{pmatrix} (y - y_j)^2 & -(x - x_j)(y - y_j) \\ \text{symm} & (x - x_j)^2 \end{pmatrix}$$

$$r_j^3 \mathbf{h}^T \mathbf{H}_j \mathbf{h} = h^2 v^2 - 2hku v + k^2 u^2 = (hv - ku)^2 \geq 0$$

f är konvex !

Optimering, IV



Hur hitta (x^*, y^*) ?

a) Newton-Raphson på
 $\mathbf{g} = 0$:

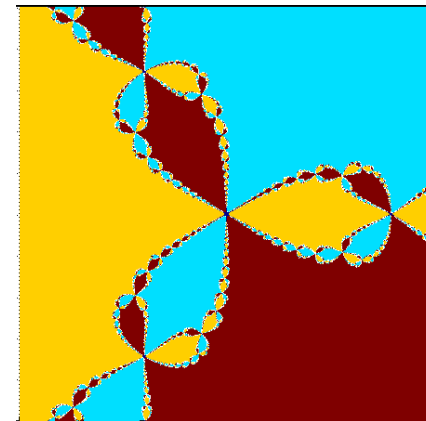
$$\mathbf{x}_{n+1} = \mathbf{x}_n - c\mathbf{H}(\mathbf{x}_n)^{-1} \mathbf{g}(\mathbf{x}_n)$$

$$d\mathbf{x}/dt = -\mathbf{H}(\mathbf{x})^{-1} \mathbf{g}(\mathbf{x})$$

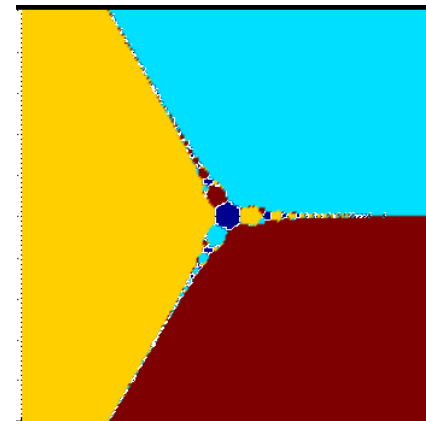
b) "Steepest descent":
 $d\mathbf{x}/dt = -\mathbf{g}(\mathbf{x})$

"Säkrare" konvergens med litet c :
Julia med

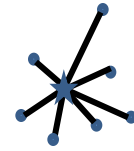
$c = 1$



$c = 0.02$



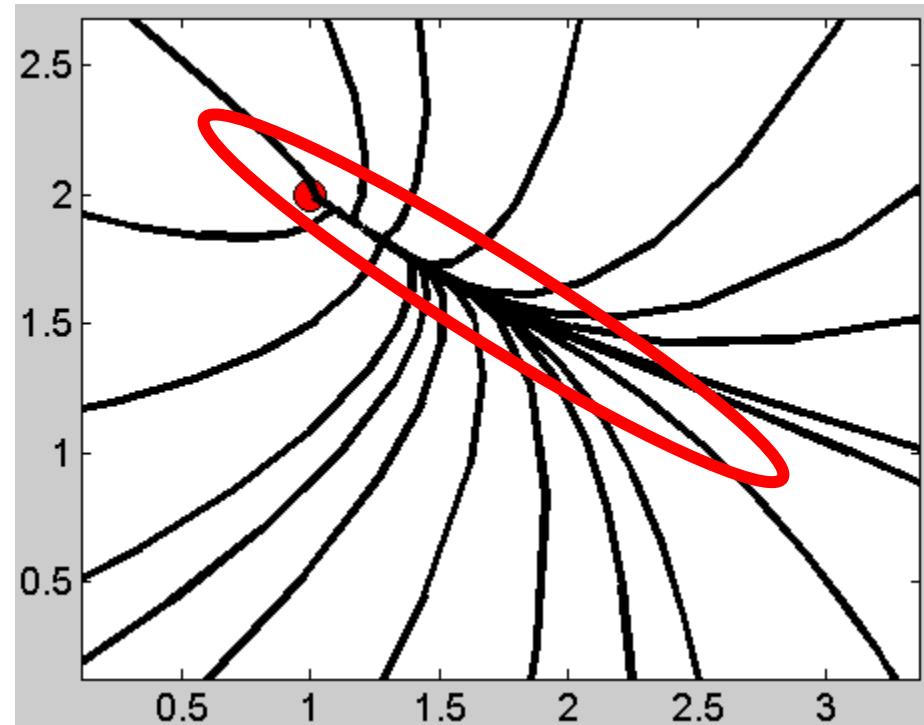
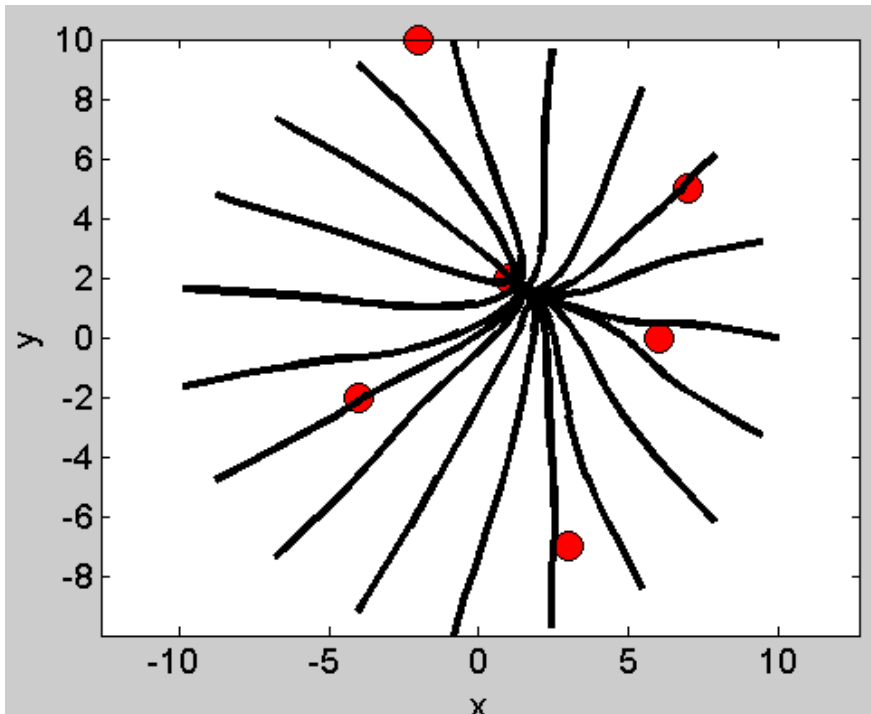
Optimering, V



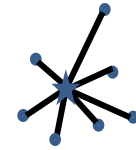
Hur hitta (x^, y^*) ?*

”Steepest descent”:

$$d\mathbf{x}/dt = -\mathbf{g}(\mathbf{x})$$



Optimering, VI



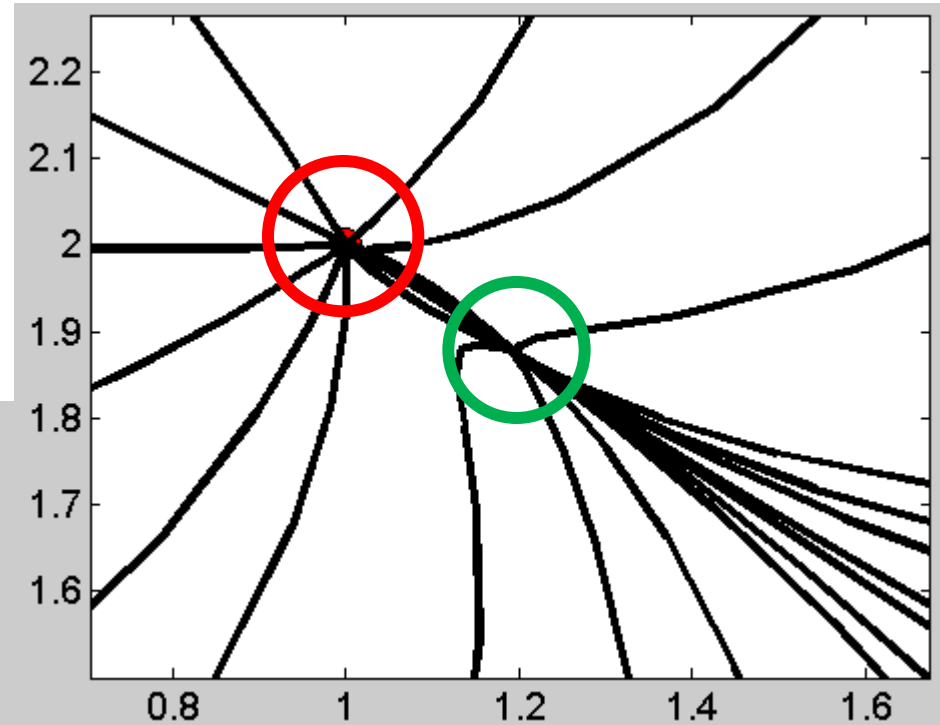
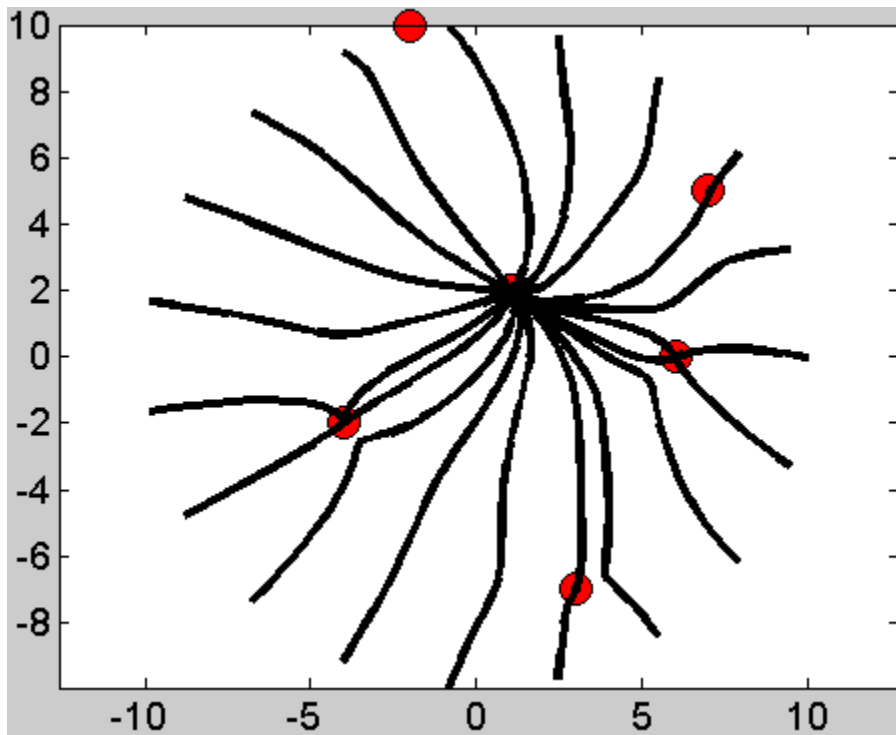
Hur hitta (x^, y^*) ?*

Newton-Raphson på

$\mathbf{g} = 0$:

$$\mathbf{x}_{n+1} = \mathbf{x}_n - c\mathbf{H}(\mathbf{x}_n)^{-1} \mathbf{g}(\mathbf{x}_n)$$

$$d\mathbf{x}/dt = -\mathbf{H}(\mathbf{x})^{-1} \mathbf{g}(\mathbf{x})$$



Vad händer med $\mathbf{H}^{-1}\mathbf{g}$
nära punkterna?

Levenberg-Marquardt:

$$(\mathbf{H} + \delta\mathbf{I})^{-1}\mathbf{g}$$

δ väljs så minsta egenvärde ökas

Optimering, VII

Problem med Steepest Descent

Testa på min $\mathbf{x}^T \mathbf{H} \mathbf{x}$, $\mathbf{g} = 2\mathbf{H}\mathbf{x}$

$$\mathbf{x}_{n+1} = \mathbf{x}_n + s \mathbf{g}(\mathbf{x}_n)$$

s med linjesökning

$$\min f(\mathbf{x} + s\mathbf{g})$$

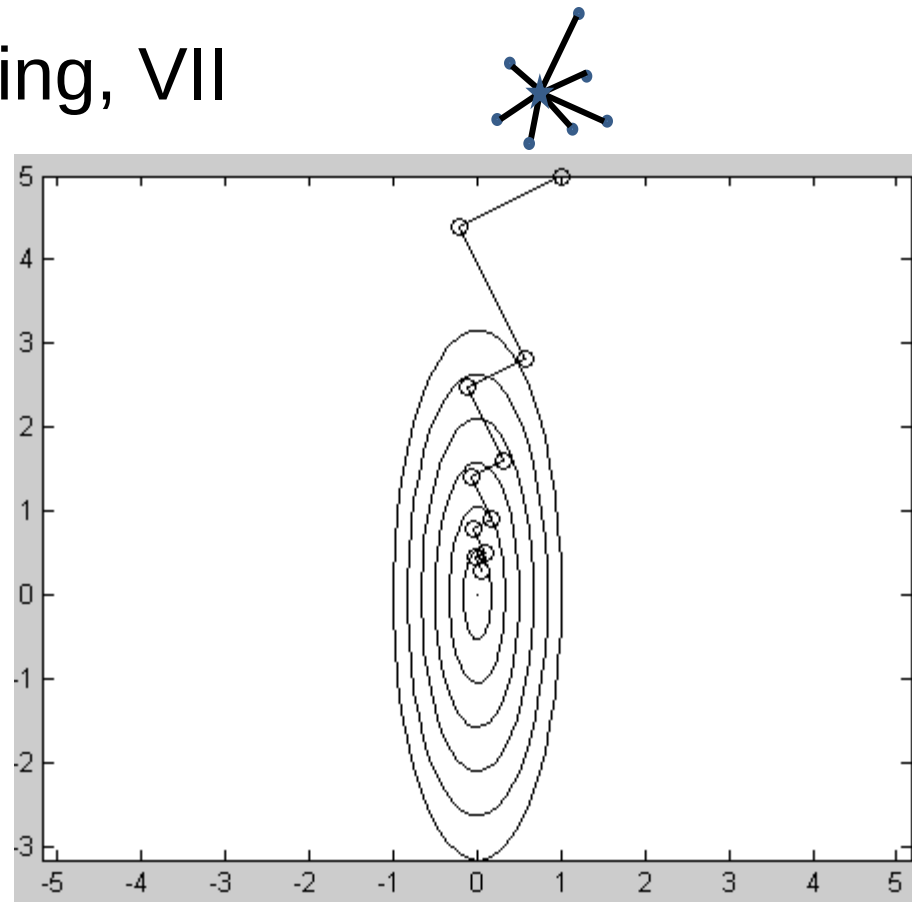
$$\mathbf{H} = \text{diag}(1, 0.1)$$

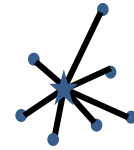
Med Newton: 1 steg till opt.

Men svårt att beräkna
andra-derivator!

Alltså: Quasi-Newton-metoder där Hessianen
approximeras, t ex

BFGS (Broyden-Fletcher-Goldfarb-Shanno)





Prova annat sätt att mäta avstånd

$$|\mathbf{u} - \mathbf{v}| = |u_1 - v_1| + |u_2 - v_2|$$

f är inte deriverbar där
 $x = x_j$ eller $y = y_j$

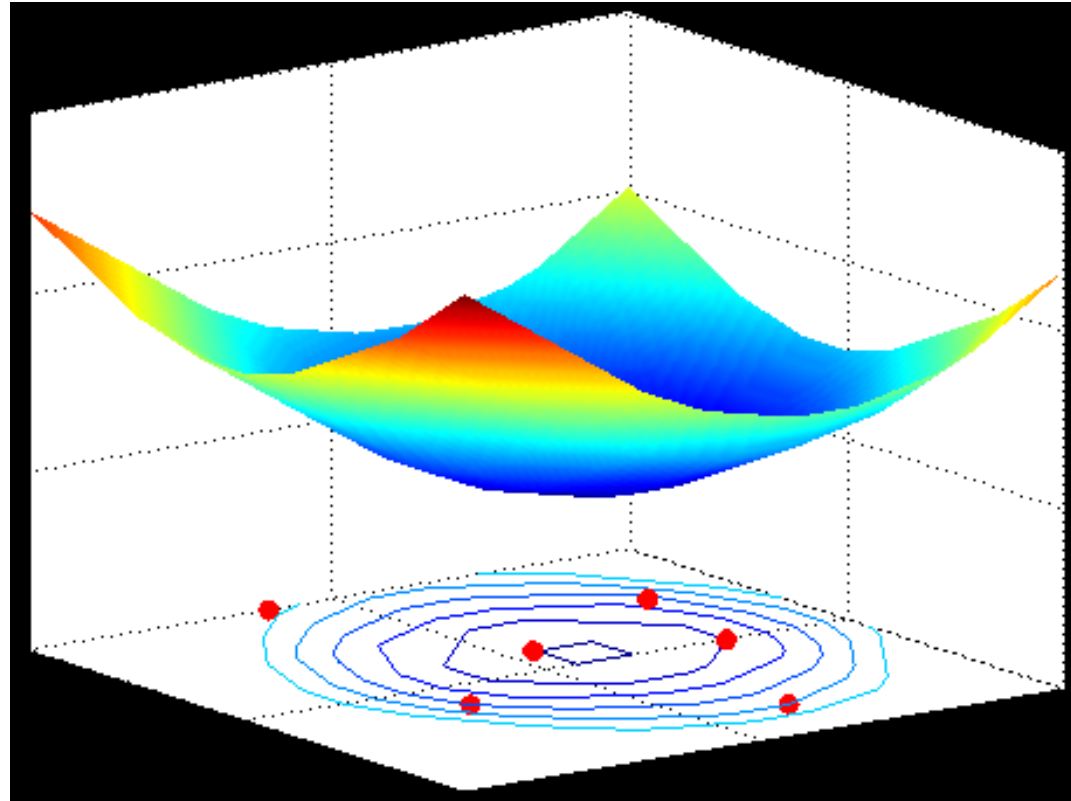
f är fortfarande konvex
 omän inte uniformt konvex.

Unik minimerare?

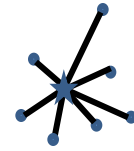
Problemet sönderfaller i ett
 för x och ett för y ,

$$\min_x m(x), m(x) = \sum_j |x - x_j|$$

m är en konvex funktion, därmed unimodal. Vi kan hitta
 lösningen med 1-D sökning (Gyllene snittet, NAM 7)



Optimering, XI



Gyllenesnittet-sökning

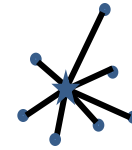
$$\min f(x), 0 < x < 1$$

- f unimodal: bara ett min. i $[0,1]$

Sökalgoritm som garanterar att min. ligger i visst intervall vars längd iterativt minskas:

- Låt intervallet till en början vara $(0,1)$ med $0 < x < 1$, så att
$$f(x) < \min(f(0), f(1))$$
- Beräkna $f(y)$, $y < x$. Intervallet med min kan nu minskas och algoritmen upprepas därefter i rätt mindre intervall:
 - min i $(0, x)$ om $f(y) < f(x)$
 - min i $(y, 1)$ om $f(y) > f(x)$
- Hur välja x & y ?

Optimering, IX



Prova annat sätt att mäta avstånd, II

$$\min_x m(x), m(x) = \sum_j |x - x_j|$$

m är min i $[1,3] = [x_3, x_4]$

Allmänt $(x_1, x_2, \dots, x_N)$:

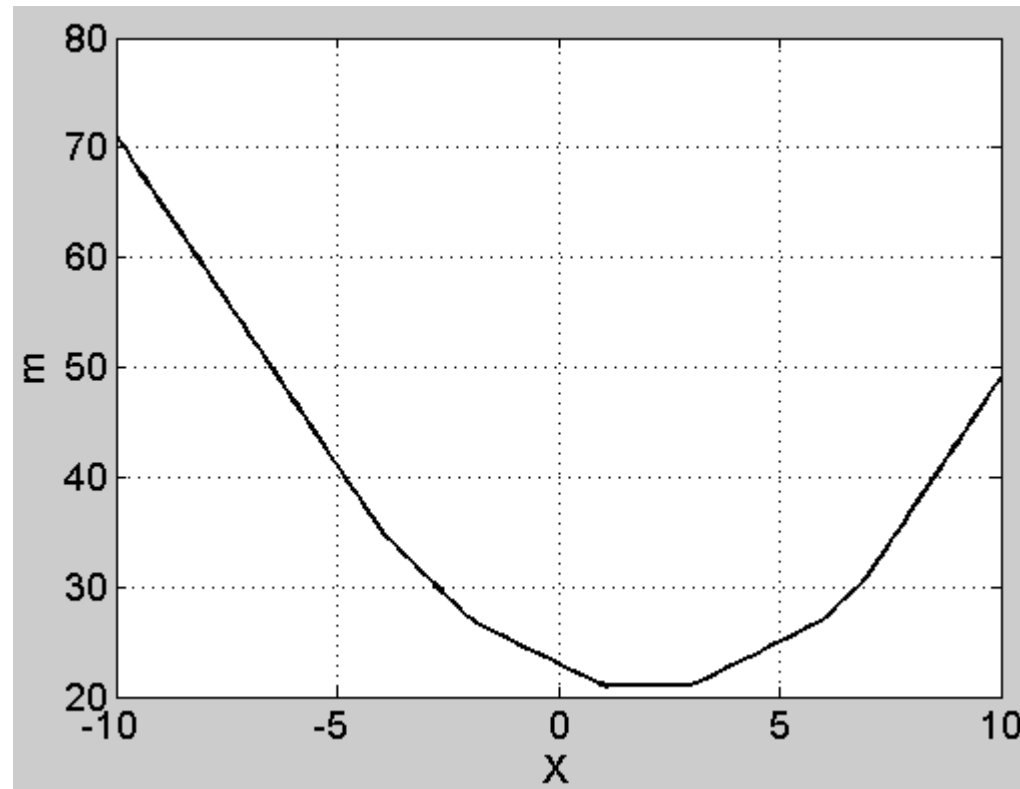
f' ökar med 2 vid varje x_j

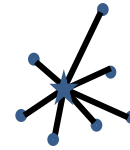
$-N$ då $x \rightarrow -\infty$,
 $+N$ då $x \rightarrow +\infty$

Medianen:

N jämnt: min i $[x_{N/2}, x_{N/2+1}]$

N udda: min i $x_{(N+1)/2}$





Gyllenesnittet-sökning, II



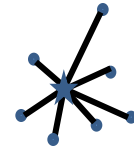
$$\frac{y}{x} = \frac{x}{1} = \frac{1-x}{1-y}$$

$$y = x^2; x - x^3 = 1 - x:$$

$$x(1-x)(1+x) = 1-x \Rightarrow x = 1, \frac{-1 \pm \sqrt{5}}{2}$$

Intervall-längden minskar med faktor $\phi = 0.68..$ per iteration

Optimering, X



Linjär-programmering (1940 ff Danzig, Wolfe, mfl.:)

$$\min \mathbf{a}^T \mathbf{x}, \mathbf{x} \in R^n$$

$$\mathbf{b}_i^T \mathbf{x} \leq 0, i = 1, 2, \dots$$

Likhetsbivillkor: $\mathbf{b}_i = -\mathbf{b}_{i+1}$

- Tillåtet (eng. feasible) område för $\mathbf{x}$ är konvext;
- Lösning i ett (eller flera) hörn där n bivillkor är uppfyllda med likhet
- Simplex-algoritm – undersök hörnen:
byt till grannhörn som ger bäst förbättring

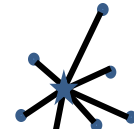
Ex.

$$\min \sum_{j=1}^{N+1} a_j \cdot x_j,$$

$$a_1 = 0, a_j = 1, j \geq 2$$

$$-x_i + X_i - x_1 \leq 0, -x_i - X_i + x_1 \leq 0, i = 1, 2, \dots, N$$

Optimering, X



Exempel, $n = 2$

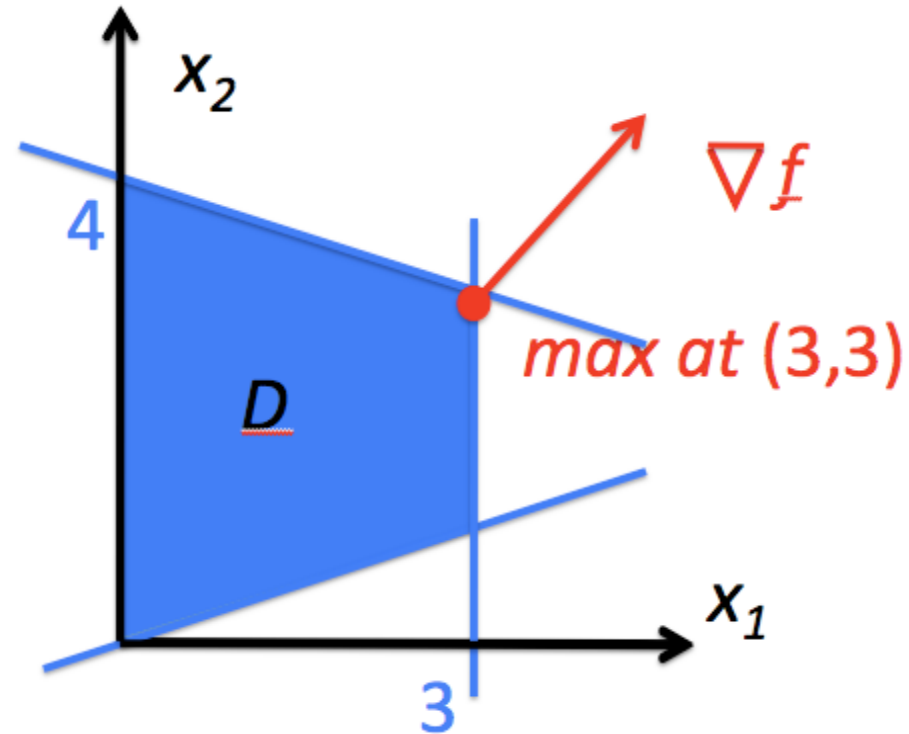
$$\max x_1 + x_2$$

$$x_1 + 3x_2 - 12 \leq 0$$

$$x_1 - 3x_2 \leq 0$$

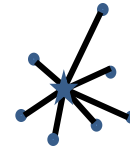
$$x_1 - 3 \leq 0$$

$$-x_1 \leq 0$$



Gradienten av objektfunktionen är $(1,1)$

Optimering, X



Problem med bivillkor (Unconstrained optimization)

- Två typer av metoder är vanliga.

(1) utveckla algoritm som inkluderar bivillkoren:

- enkla gränser enkelt,
- SQP, Sequential Quadratic Programming
- Projicerad gradient, reducerad gradient, ...

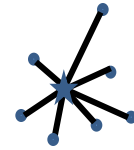
(2) Transformera problemet till problem utan bivillkor

a) Ex. $x > 0$, inför x som y^2 .

b) Strafffunktionsteknik. Addera strafffunktion till $f(x)$ så att bivillkoren uppfylls automatiskt vid maximering.

$$\begin{aligned} \max_{x \in D} f(x) &\rightarrow \max_x f(x) + \varepsilon g_D(x) \\ g_D(x) &\downarrow -\infty \text{ då } x \rightarrow \partial D, x \in D \end{aligned}$$

Optimering, X

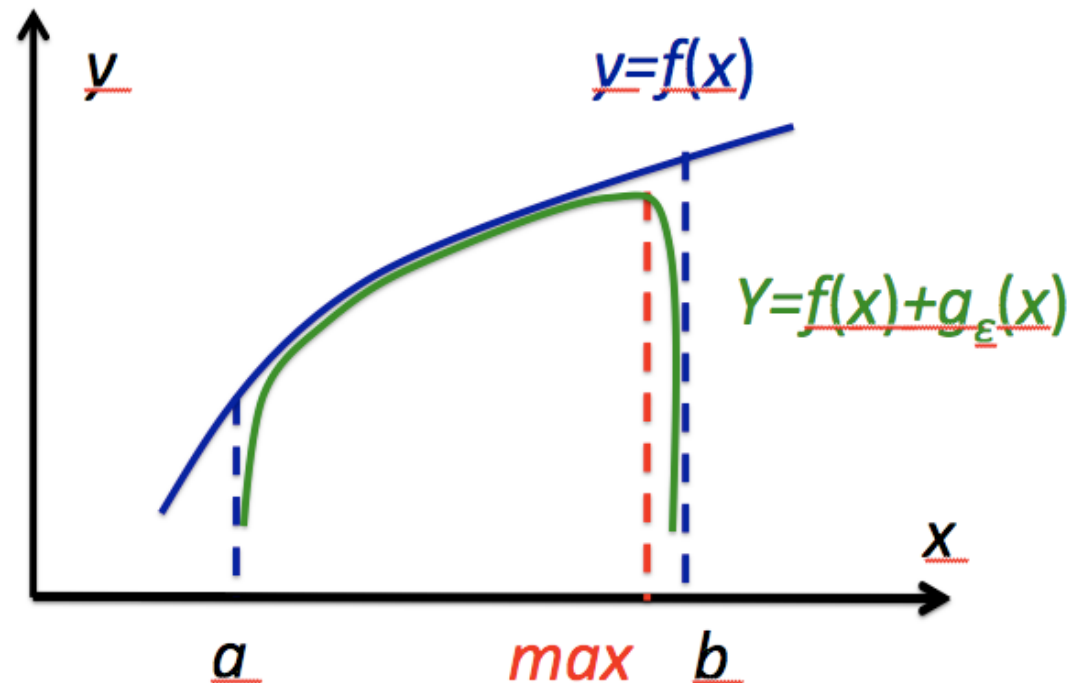


Exempel, strafffunktionsteknik.

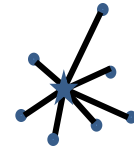
- Starta med x_0 i (a,b) och finn x som ger max
- Använd detta x -värde som start för ny maximering med mindre ε

$$\max_{a < x < b} f(x) \rightarrow \max_x f(x) + g_\varepsilon(x)$$

$$g_\varepsilon(x) = \frac{\varepsilon}{(x-a)(b-x)}$$

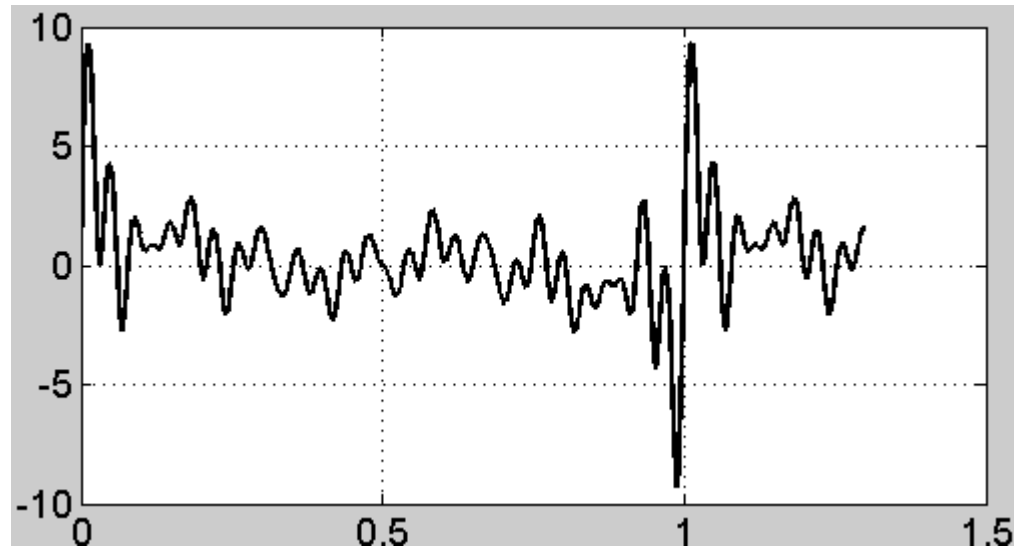


Optimering, X



Hur hitta globalt optimum?

- för vissa problem vet man att det finns 1 stationär punkt
- gradient-metoder närsynta
- Ex. 30 termer i Fourier-serie



Prova olika startvärden

Genetisk algoritm

Simulated annealing / Metropolis