Three examples

- Ex1: The need of line search
  - The Rosenbrock function

$$f(x,y) = 100(y - x^2)^2 + (1 - x)^2$$

- Ex2: The modified Newton's direction

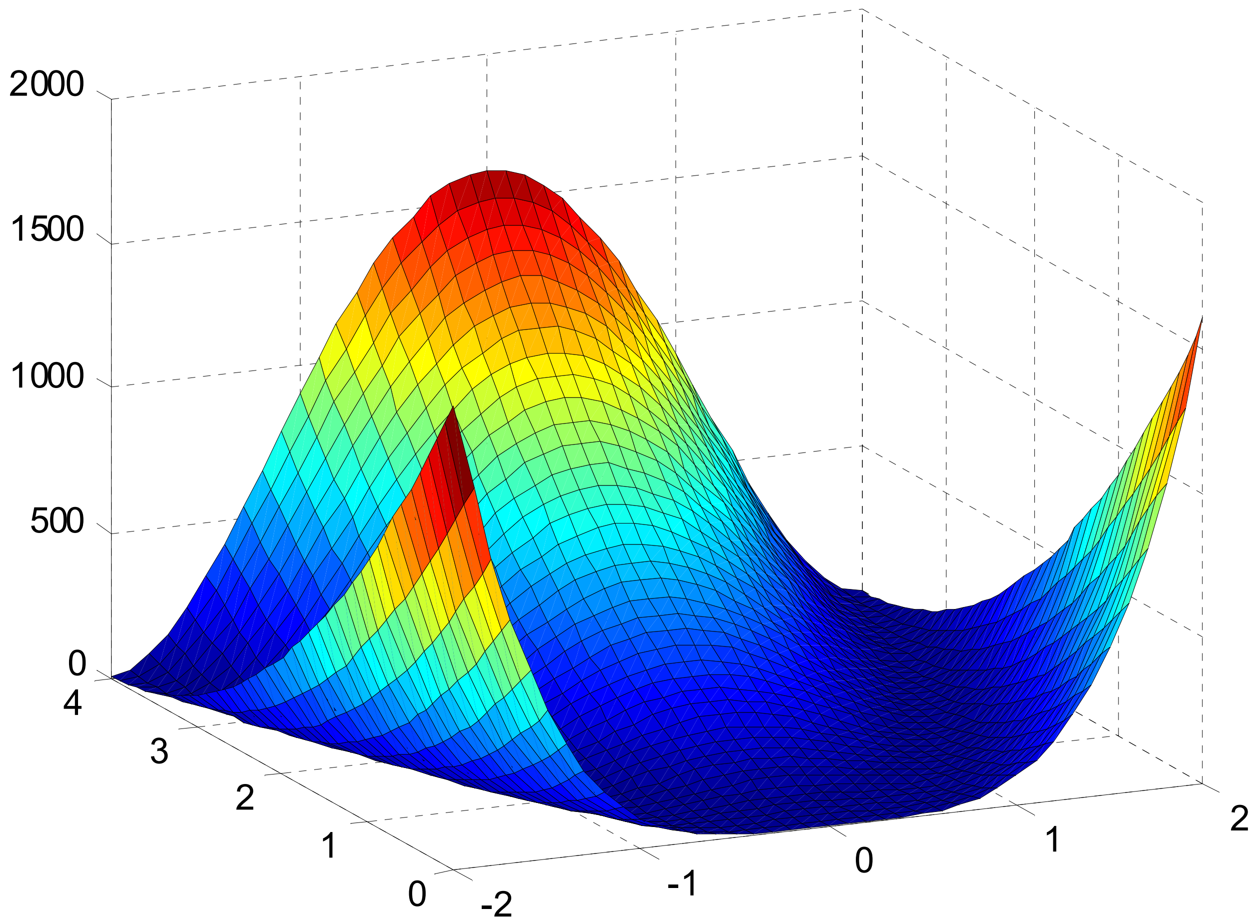
$$\min_{0 \le x, y \le 4} f(x,y) = x^2 - y^2$$

- Ex3: The descent direction

$$f(x) = x^2$$

The Rosenbrock function

$$f(x, y) = 100(y - x^2)^2 + (1 - x)^2$$



Gradient and Hessian

$$f(x,y) = 100(y - x^2)^2 + (1 - x)^2$$

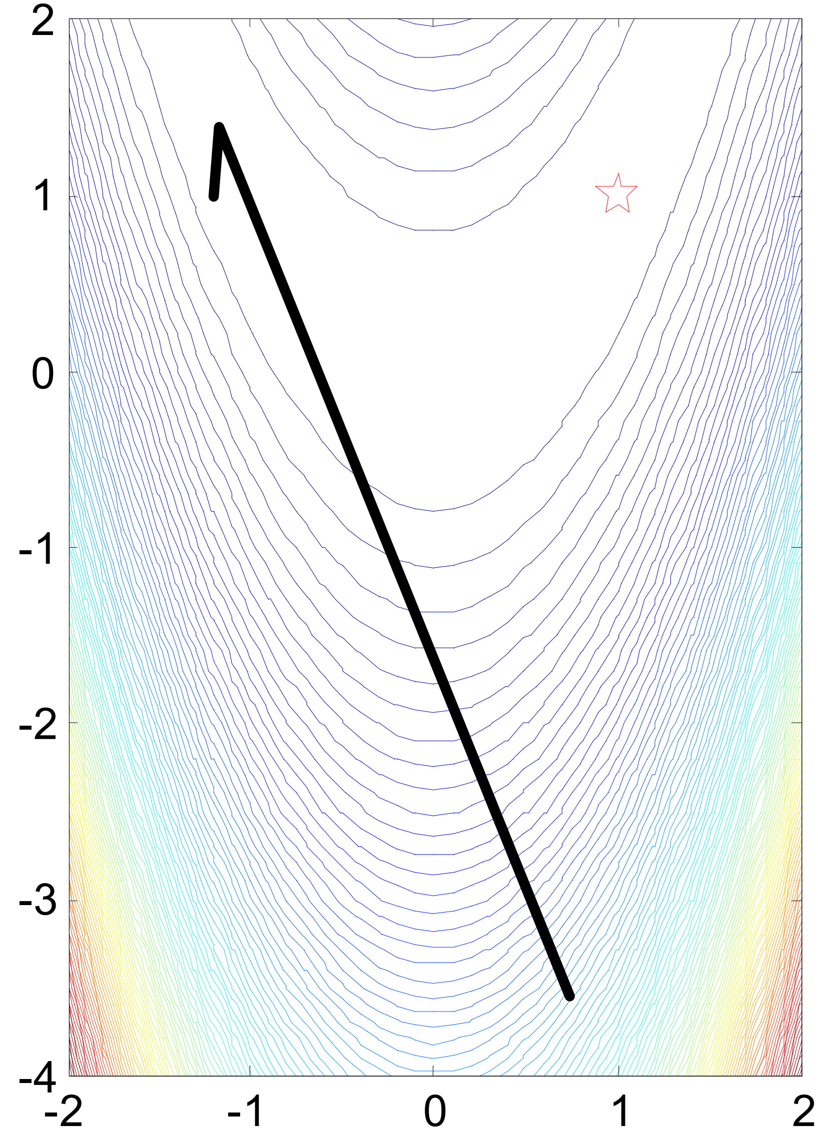
$$\nabla f(x,y) = \begin{pmatrix} -400(y - x^2)x - 2(1 - x) \\ 200(y - x^2) \end{pmatrix}$$

$$\nabla^2 f(x,y) = \begin{pmatrix} -400(y - 3x^2) + 2 & -400x \\ -400x & 200 \end{pmatrix}$$

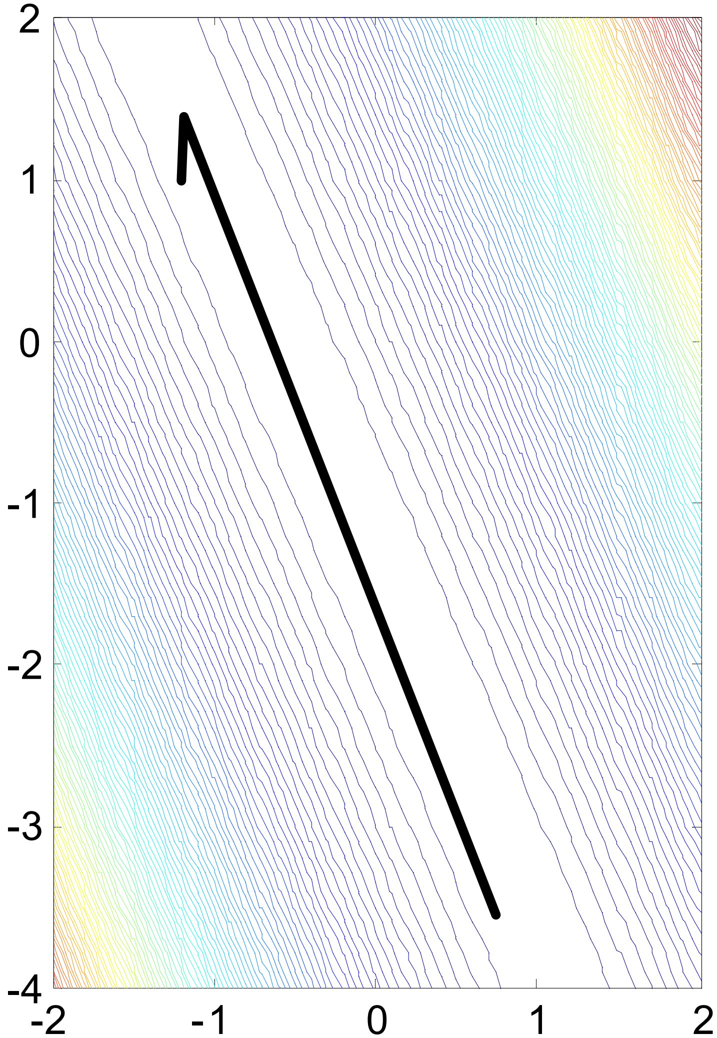
Minimizer is at (1,1)

Newton's direction

- Let $x_0 = (-1.2, 1.0)$
- Compute Newton's direction
- $x_1 = x_0 + p_1 = (-1.17, 1.38)$
- $x_2 = x_1 + p_2 = (0.74, -3.56)$

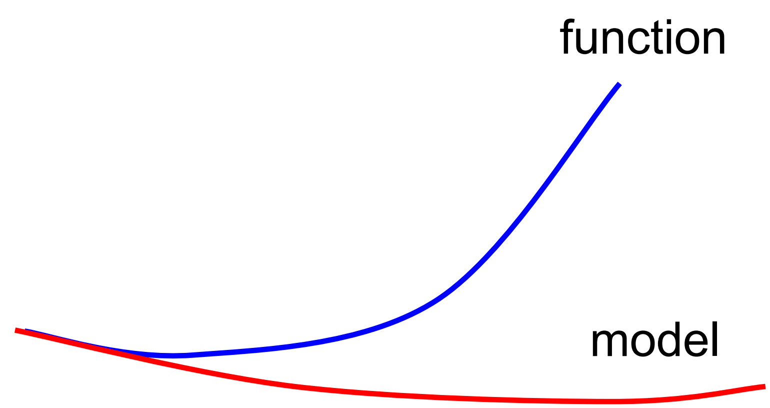


Model function

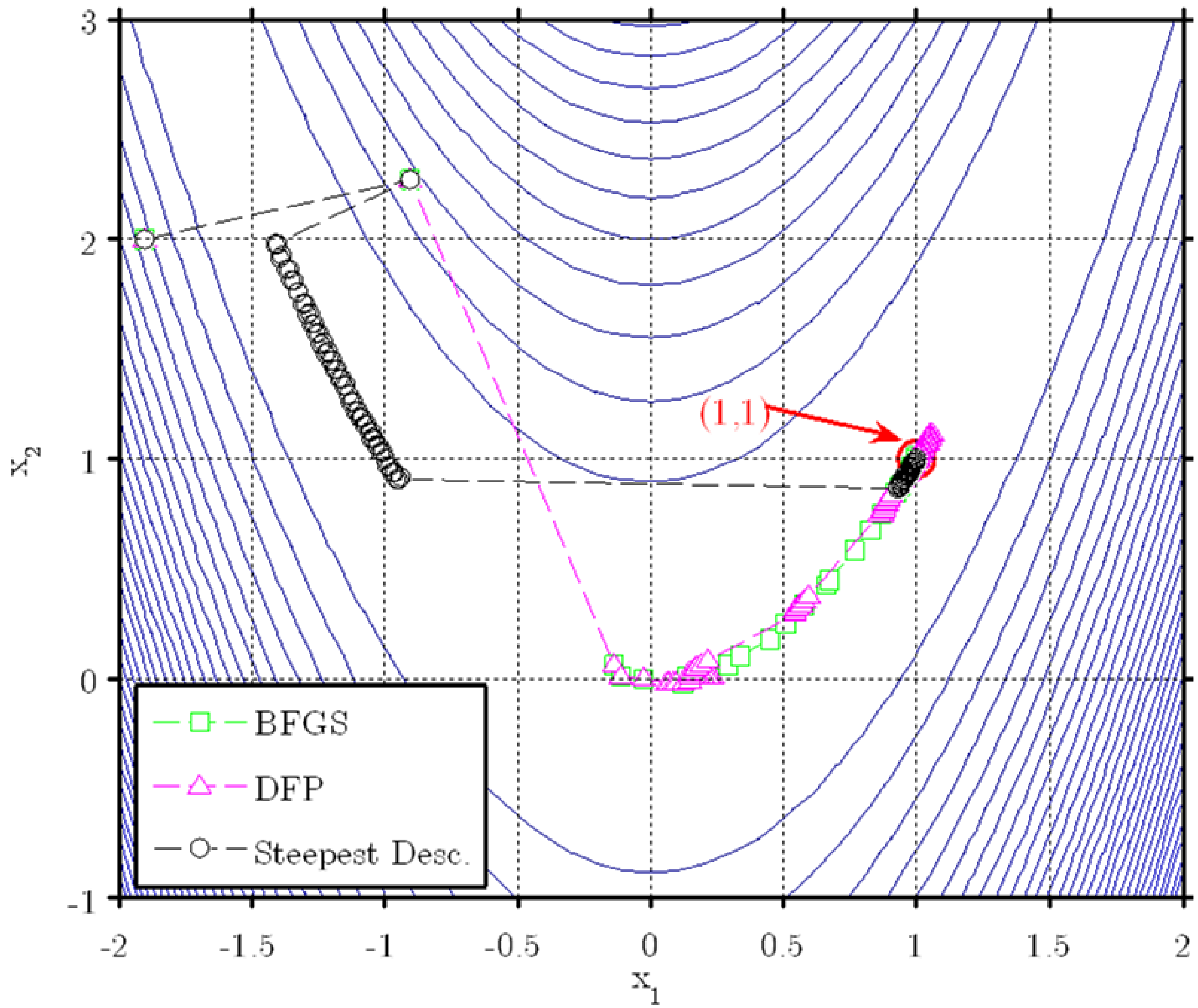


$$\nabla^2 f(x_2) =$$

$$\begin{pmatrix} 1107.27 & 470.11 \\ 470.11 & 200 \end{pmatrix}$$



http://www.rpi.edu/~vanfrl/rosenbrock_minim.gif



Example 2

$$\min_{0 \le x, y \le 4} f(x,y) = x^2 - y^2$$

- – A constrained optimization problem
- – Minimizer is at (x,y)=(0,4)

$$\nabla f = \begin{pmatrix} 2x \\ -2y \end{pmatrix}, \nabla^2 f = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$$

- Newton's direction

$$p^{\mathrm{N}} = -(\nabla^2 f)\nabla f = \begin{pmatrix} -x \\ -y \end{pmatrix}$$

Modified Newton's direction

- At (x,y)=(1,2), f(x,y) = 1 − 4= −3.
  - The Newton's direction $p^N = (-1; -2)$
  - $(x^+,y^+) = (1-1, 1-1) = (0,0)$ and $f(x^+,y^+) = 0$.
- Modified the Hessian,

$$\nabla^2 f = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

- Modified Newton's direction $\tilde{p}^{\mathrm{N}} = \begin{pmatrix} -x \\ y \end{pmatrix}$
  - $(x^+,y^+) = (1-1, 2+2) = (0,4)$ and $f(x^+,y^+) = -16$

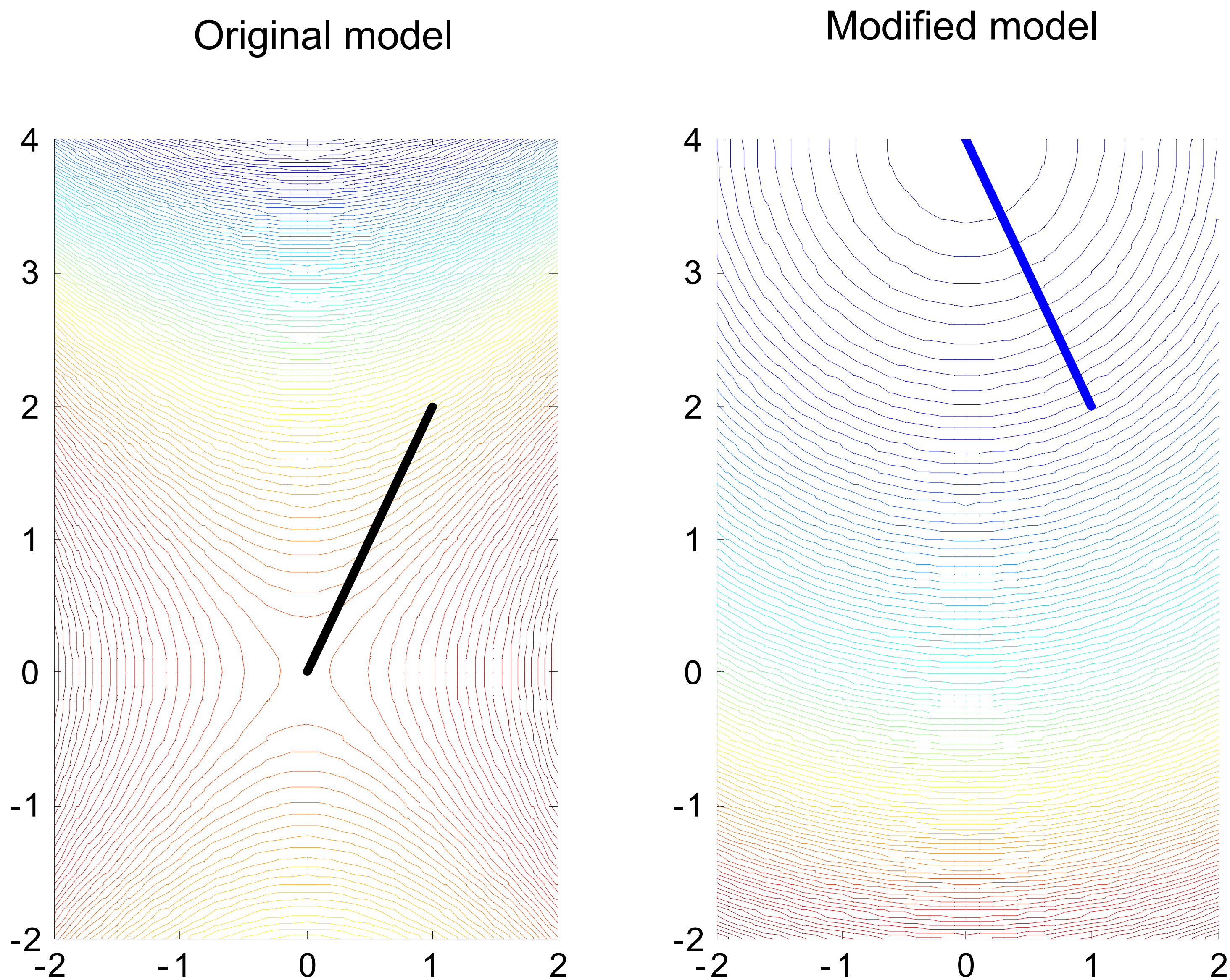
The function of modified Newton

At (x,y)=(1,2), the original model function

$$\begin{aligned} f(z) &\approx f(x_0) + \nabla f(x_0)^T (z - x_0) + \\ & \frac{1}{2}(z - x_0)^T \nabla^2 f(x_0)(z - x_0) = x^2 - y^2 \end{aligned}$$

- The modified model

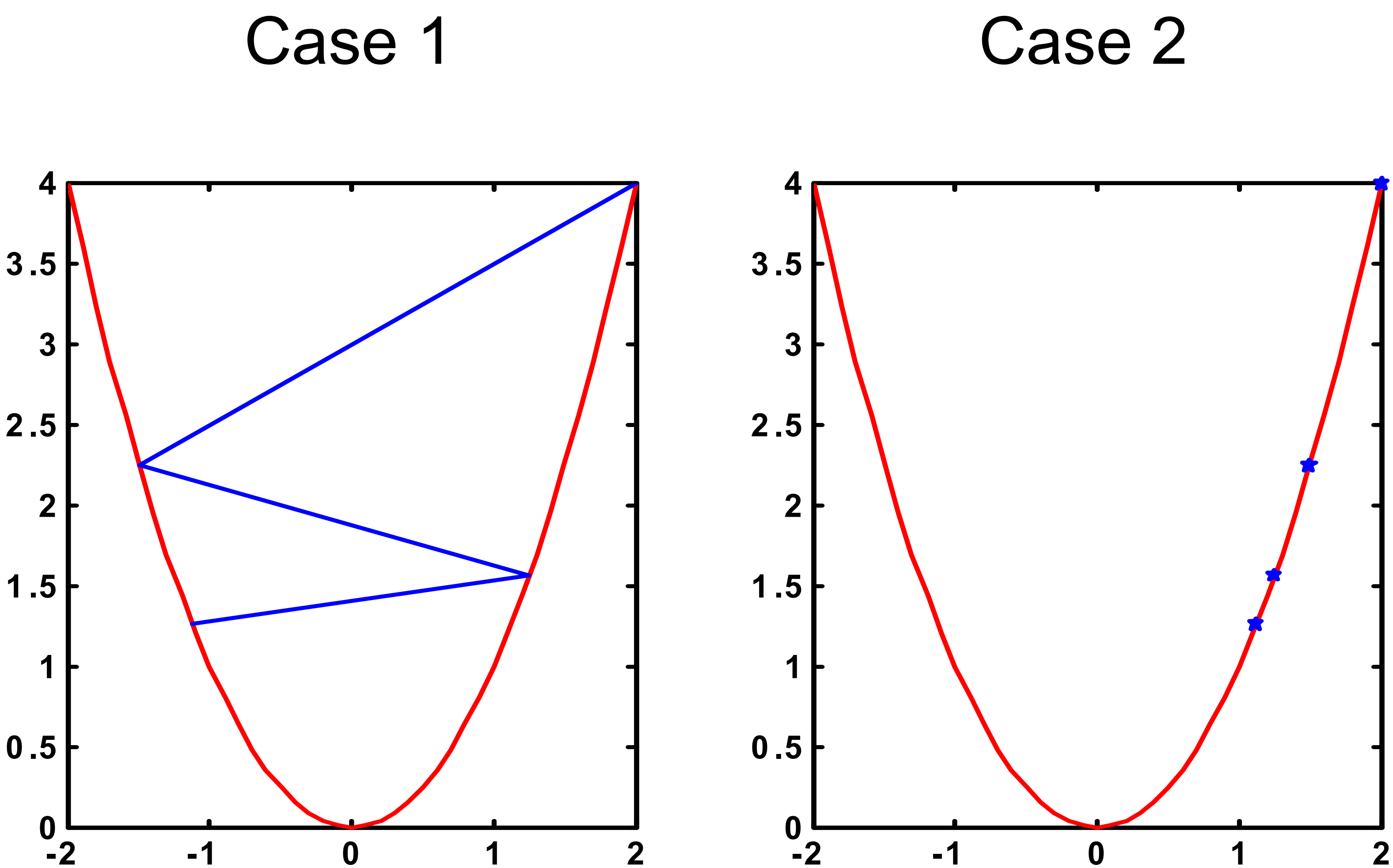
$$\begin{aligned} f(z) &\approx f(x_0) + \nabla f(x_0)^T (z - x^0) + \frac{1}{2}(z - x_0)\hat{H}(z - x_0) \\ &= x^2 + y^2 - 8y + 8 \end{aligned}$$



Original model
Modified model
4
3
2
1
0
-1
-2
-2
-1
0
1
2

Example 3

- $f(x) = x^2$, $x_0 = 2$,
  - Case 1,
  - Step $p_k = (-1)^k$
  - Step length $\alpha_k = 2+3*2^{-(k+1)}$
  - x_k={2,-3/2,5/4,…} $\to \pm 1$
  - $f(x_k) \to 1$
  - Case 2,
  - Step $p_k = -1$
  - Step length $\alpha_k = 2+3*2^{-(k+1)}$
  - x_k={2,3/2,5/4,…} $\to 1$
  - $f(x_k) \to 1$



Case 1
Case 2
4
3.5
3
2.5
2
1.5
1
0.5
0
-2
-1
0
1
2