MATH529 – Fundamentals of Optimization

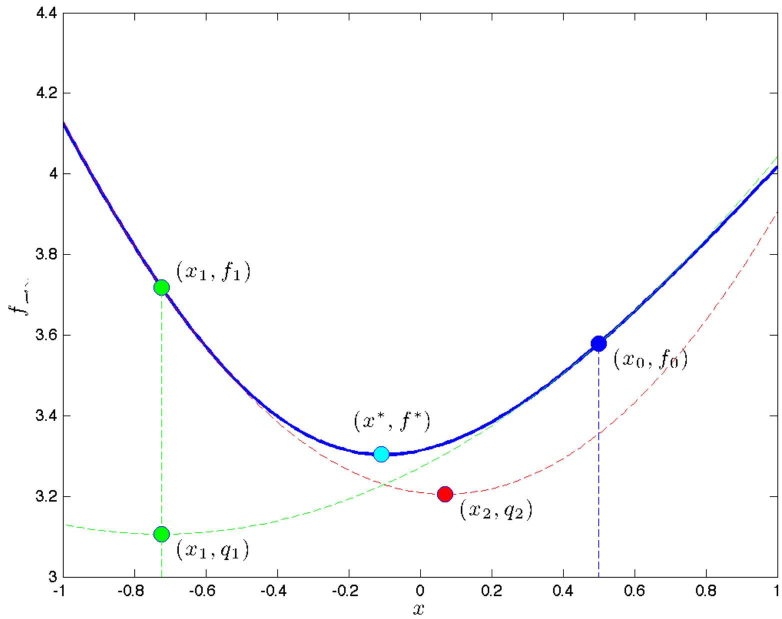
Unconstrained Optimization V

Marco A. Montes de Oca

Mathematical Sciences, University of Delaware, USA

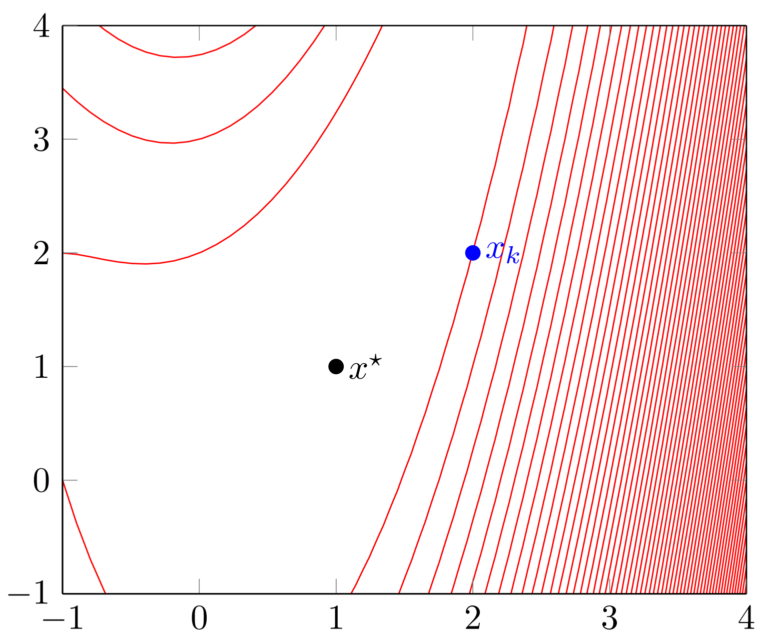
Recap

Newton's Method as an Optimization Algorithm

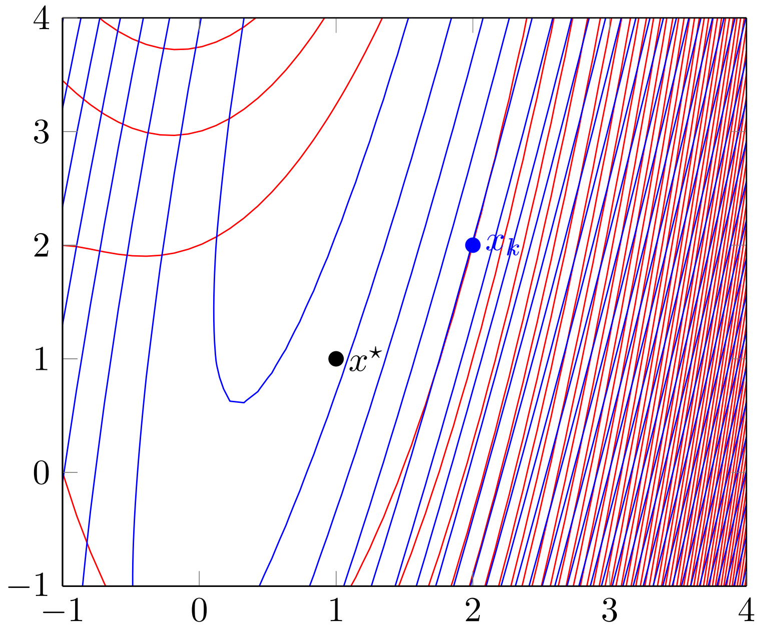


Taken from `http://inst.eecs.berkeley.edu/~ee127a/book/login/l_cp_algs.html`

Newton's Method as an Optimization Algorithm



Newton's Method as an Optimization Algorithm



Newton's Method as a Line Search Method

As a root finder:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \left(\nabla \mathbf{g}(\mathbf{x}_k)\right)^{-1} \mathbf{g}(\mathbf{x}_k)$$

As an optimization algorithm:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \left(Hf(\mathbf{x}_k)\right)^{-1} \nabla f(\mathbf{x}_k)$$

In general:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{p}_k$$

Newton's Method as a Line Search Method

In this case:

$$\alpha_k = 1$$

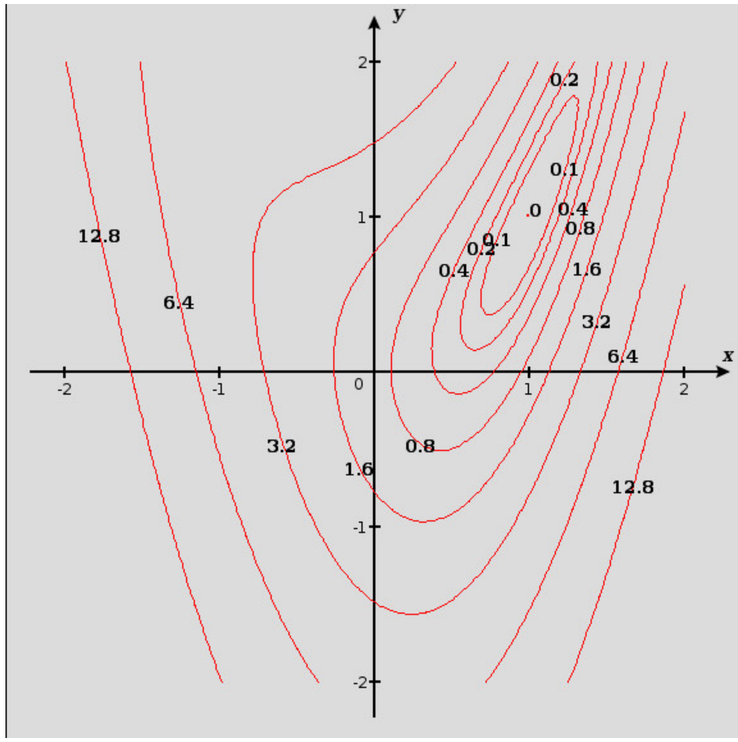
and

$$\mathbf{p}_k = -\left(Hf(\mathbf{x}_k)\right)^{-1} \nabla f(\mathbf{x}_k)$$

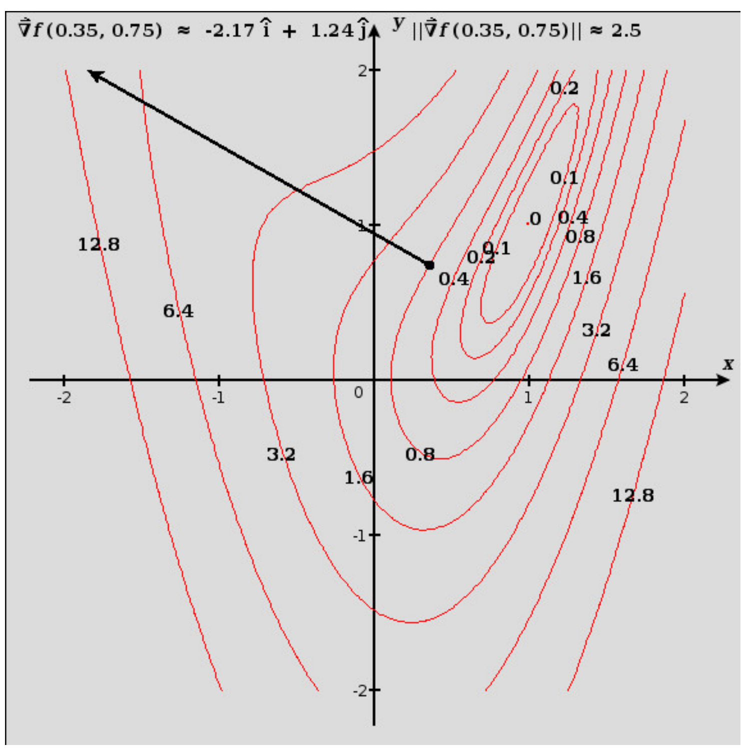
Newton direction

Another line search method: Steepest descent

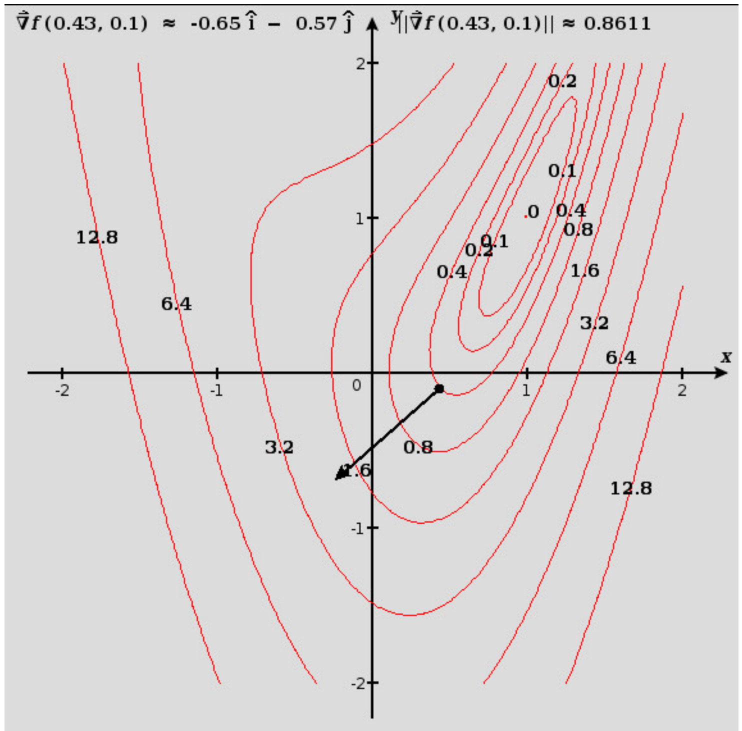
Contour plot of $f(\mathbf{x}) = (x_2 - x_1^2)^2 + (1 - x_1)^2$



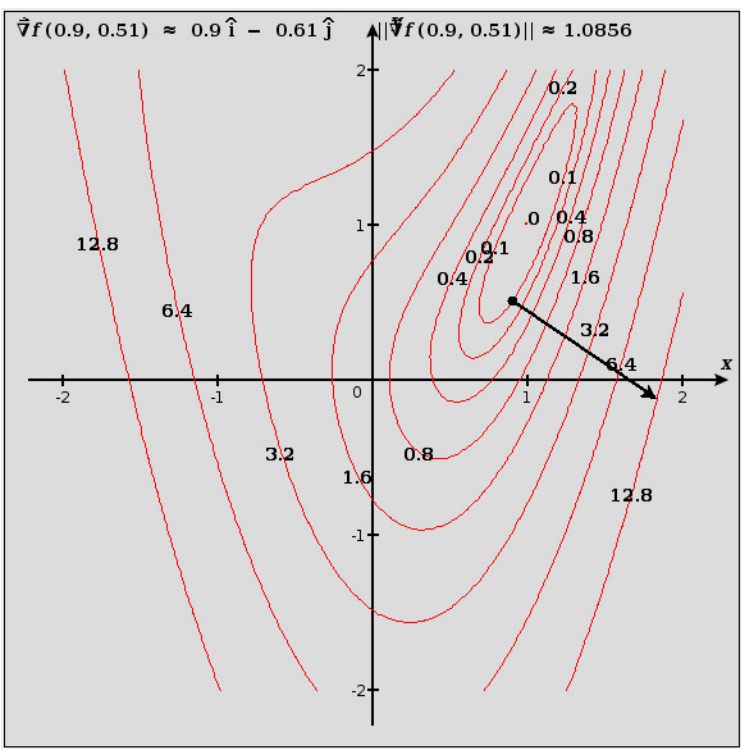
Contour plot of $f(\mathbf{x}) = (x_2 - x_1^2)^2 + (1 - x_1)^2$



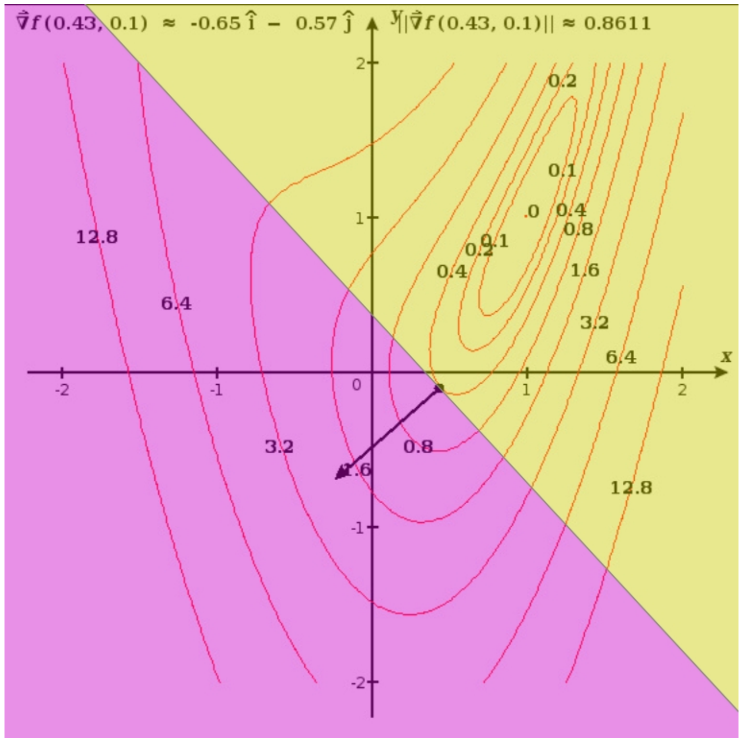
Contour plot of $f(\mathbf{x}) = (x_2 - x_1^2)^2 + (1 - x_1)^2$



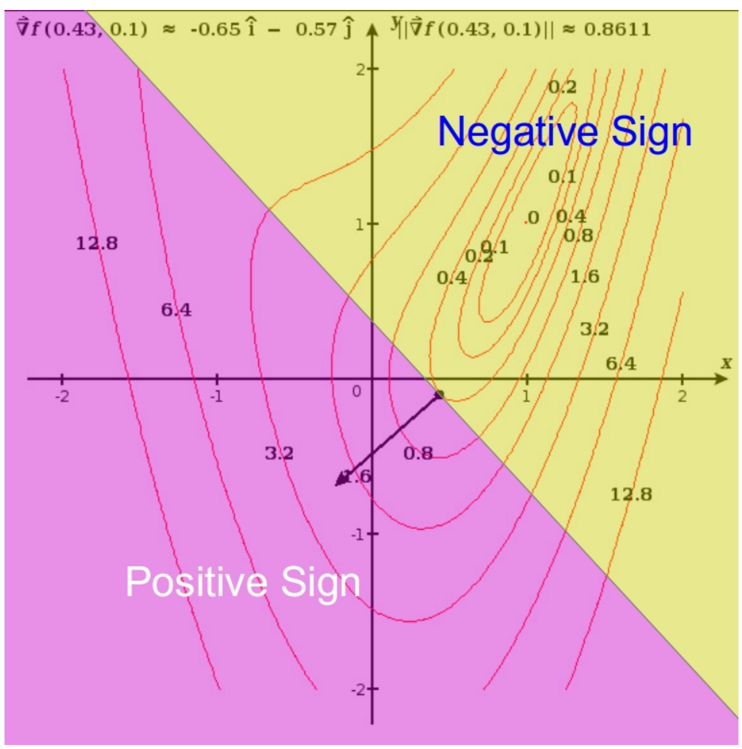
Contour plot of $f(\mathbf{x}) = (x_2 - x_1^2)^2 + (1 - x_1)^2$



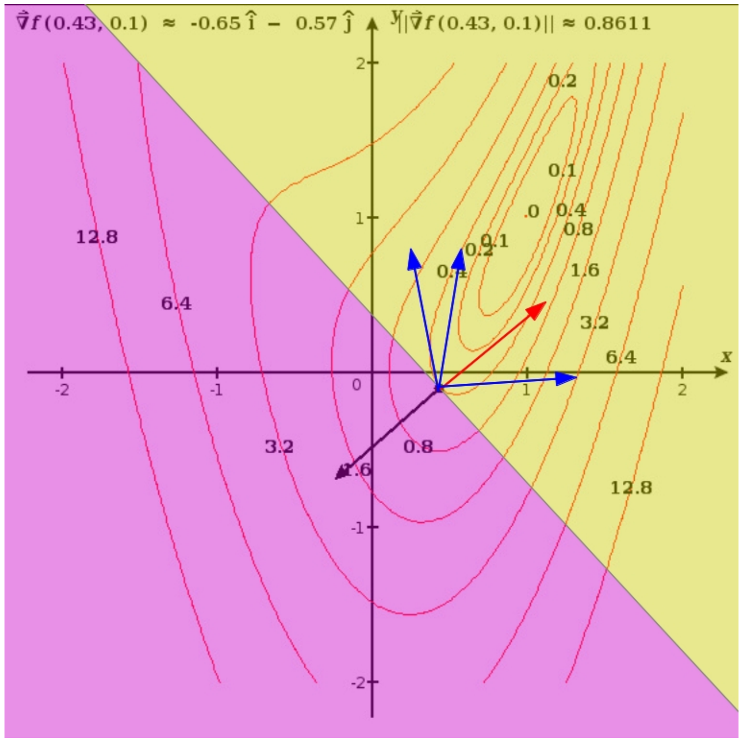
Sign of Directional Derivatives



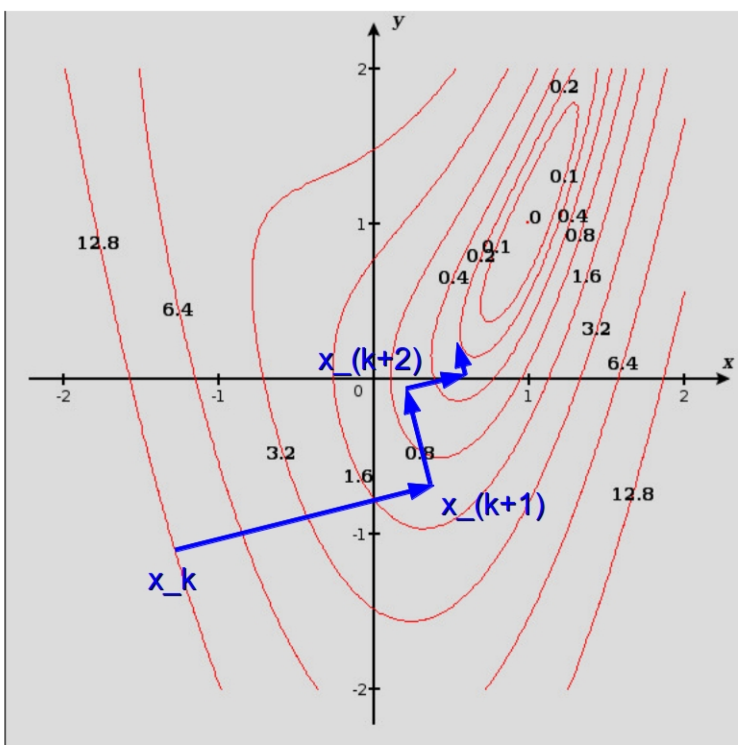
Sign of Directional Derivatives



Descent Search Directions



Line Search Algorithms



Steepest Descent Direction

Let's $\mathbf{p}_k$ denote the search direction chosen at step k, that is, $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{p}_k$. Then, according to Taylor's formula, we have:

$f(\mathbf{x}_{k+1}) = f(\mathbf{x}_k + \alpha_k \mathbf{p}_k) =$
$f(\mathbf{x}_k) + \alpha_k \nabla f(\mathbf{x}_k)^T \mathbf{p}_k + \frac{\alpha_k^2}{2} \mathbf{p}_k^T Hf(\mathbf{x}_k + t\mathbf{p}_k)\mathbf{p}_k$
for some $t \in (0, \alpha_k)$.

Then,

$\frac{\mathrm{d}f(\mathbf{x}_k + \alpha_k \mathbf{p}_k)}{\mathrm{d}\alpha_k} = \nabla f(\mathbf{x}_k)^T \mathbf{p}_k + \alpha_k \mathbf{p}_k^T Hf(\mathbf{x}_k + t\mathbf{p}_k)\mathbf{p}_k$.
When $\alpha_k = 0$, that is, the rate of change of f at $\mathbf{x}_k$ in the direction of $\mathbf{p}_k$ is: $\nabla f(\mathbf{x}_k)^T \mathbf{p}_k = ||\nabla f(\mathbf{x}_k)|| ||\mathbf{p}_k|| \cos\theta$, where θ is the angle between $\mathbf{p}_k$ and $\nabla f(\mathbf{x}_k)$.

Thus, the direction of most rapid decrease at $\mathbf{x}_k$ is given by:$\mathbf{p}_k = -\frac{\nabla f(\mathbf{x}_k)}{||\nabla f(\mathbf{x}_k)||}$.

Basic search direction selection

Steepest descent direction:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \alpha_k \frac{\nabla f(\mathbf{x}_k)}{||\nabla f(\mathbf{x}_k)||}, \text{ or simply } \mathbf{x}_{k+1} = \mathbf{x}_k - \alpha_k \nabla f(\mathbf{x}_k)$$

Newton's direction:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - (Hf(\mathbf{x}_k))^{-1} \nabla f(\mathbf{x}_k),$$

Micro-Lab