7 | Minimization

Methods

The goal is to minimize

$$F(\mathbf{x}) = F(x_1, x_2, \ldots, x_k), \text{ where } \mathbf{x} = [x_1, x_2, \ldots, x_k]$$

This is equivalent to solving

$$\nabla F(x_1, x_2, \ldots, x_n) = \begin{bmatrix} F_1(x_1, x_2, \ldots, x_n) \\ F_2(x_1, x_2, \ldots, x_n) \\ \vdots \\ F_n(x_1, x_2, \ldots, x_n) \end{bmatrix} = \mathbf{0}$$

Definitions

The gradient:

$$\nabla F(\mathbf{x}) = \nabla F(x_1, x_2, \ldots, x_n) = \begin{bmatrix} \frac{\partial F}{\partial x_1} & \frac{\partial F}{\partial x_2} & \cdots & \frac{\partial F}{\partial x_k} \end{bmatrix}$$

The Hessian matrix:

$$HF(\mathbf{x}) = HF(x_1, x_2, \ldots, x_n) = \begin{bmatrix} \frac{\partial^2 F}{\partial x_1 \partial x_1} & \frac{\partial^2 F}{\partial x_2 \partial x_1} & \cdots & \frac{\partial^2 F}{\partial x_k \partial x_1} \\ \frac{\partial^2 F}{\partial x_2 \partial x_1} & \frac{\partial^2 F}{\partial x_2 \partial x_1} & \cdots & \frac{\partial^2 F}{\partial x_k \partial x_2} \\ \vdots & . & \cdots & \vdots \\ \frac{\partial^2 F}{\partial x_1 \partial x_k} & \frac{\partial^2 F}{\partial x_2 \partial x_k} & \cdots & \frac{\partial^2 F}{\partial x_k \partial x_k} \end{bmatrix}$$

Newton's Method

It comes from the Taylor's formula in higher dimension. We find a root of x,close to your initial guess x^0 , such that $\nabla F(\mathrm{x}) = 0$ which correspond to a minimum of $F(\mathrm{x})$. You start with the function ∇F, the intial guess x^0, a tolerance ε, and a max number of iteration $ITMAX$.

1. Calculate the next iteration (pick a small constant α)

$$\mathrm{x}^{n+1} = \mathrm{x}^n - \alpha \left[HF(\mathrm{x}^n)\right]^{-1} \nabla F(\mathrm{x}^n)$$

2. If convergence is satisfactory, i.e. $||\nabla F(\mathrm{x}^{n+1})|| < \varepsilon$, or $||\mathrm{x}^{n+1} - \mathrm{x}^n|| < \varepsilon$, or iteration is $ITMAX$, return x^{n+1} and stop.

Gradient Descent

We find the minimum of $F(\mathbf{x})$. You start with the function ∇F, the intial guess $\mathbf{x}^0$, a tolerance ε, and a max number of iteration $ITMAX$.

1. Calculate the next iteration (pick a small constant α)

$$\mathbf{x}^{n+1} = x^n - \alpha \nabla F(\mathbf{x}^n)$$

2. If convergence is satisfactory, i.e. $||\nabla F(\mathbf{x}^{n+1})|| < \varepsilon$, or $||\mathbf{x}^{n+1} - \mathbf{x}^n|| < \varepsilon$, or iteration is $ITMAX$, return $\mathbf{x}^{n+1}$ and stop.