Astronomical Data Analysis II: 27-28 August 2002, Waikoloa, Hawaii, USA

J.-L Starck Fionn Murtagh


\[ \min_{x \in \mathbb{R}^n} f(x) = \min_{x \in \mathbb{R}^n} g(x) \]

where f(x) is the objective function, which we want to minimize, and g(x) is the constraint function. The solution to this problem is found by analyzing the gradient of f(x) and finding the point where it is zero. The gradient of f(x) is given by:

\[ \nabla f(x) = \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{bmatrix} \]

The gradient is zero when:

\[ \nabla f(x) = \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \]

Solving this system of equations yields the minimum point of f(x). This method is called the gradient descent method and is one of the most popular methods for solving optimization problems in machine learning. It is a simple and effective method, but it can be slow to converge. More advanced methods, such as Newton's method, can be used to improve the convergence rate. In the next section, we will discuss some advanced optimization algorithms that can be used to solve optimization problems in machine learning. In summary, the gradient descent method is a simple and effective method for solving optimization problems in machine learning. It is a popular method, but it can be slow to converge. More advanced methods, such as Newton's method, can be used to improve the convergence rate. In the next section, we will discuss some advanced optimization algorithms that can be used to solve optimization problems in machine learning.