Machine Learning for Engineering and Science Applications

Brief Introduction to Constrained Optimization

Constrained Optimization

- The general constrained optimization task is to maximize or minimize a function $f(\mathbf{x})$ by varying $\mathbf{x}$ *given certain constraints on x*
  - For example, find minimum of $f(x_1, x_2, x_3) = x_1^2 + 2x_2^2 + x_3^2$, where $\|x\|_2 \geq 1$

- Very common to encounter this in engineering practice
  - For example, designing the fastest vehicle with a constraint on fuel efficiency

- All constraints can be converted to two types of constraints
  - Equality constraints – e.g $\text{Minimize f}(x_1, x_2, x_3)$ subject to $x_1 + x_2 + x_3 = 1$
  - Inequality constraints – e.g. $\text{Minimize f}(x_1, x_2, x_3)$ subject to $x_1 + x_2 + x_3 < 1$

Equallity

$g(\vec{x})$

$\rightarrow x_1 + x_2 + x_3 - 1 = 0$

$x_1 + x_2 + x_3 - 1 < 0$

$h(x)$

- Canonical form – All optimization problems can be written as

Feasible points

Minimize $f(\mathbf{x})$ subject to the constraint that $\mathbf{x} \in \mathbb{S}$.

$$\mathbb{S} = \{\boldsymbol{x} \mid \forall i, g^{(i)}(x) = 0 \text{ and } \forall j, h^{(j)}(x) \leq 0\}$$

Multiple Equality constraints Multiple Inequality constraints

Generalized Lagrange function

$f(x)$

- The **constrained optimization problem** requires us to minimize the function while ensuring that the point discovered belongs to the feasible set.

- There are several techniques that achieve this but it is, in general, a difficult problem.

- A very common approach is to define a new function called the generalized Lagrangian

Original fn

$$L(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\alpha}) = \boxed{f(\boldsymbol{x})} + \sum_{i} \lambda_i g^{i}(\boldsymbol{x}) + \sum_{j} \alpha_j h^{(j)}(\boldsymbol{x})$$

- Then, the constrained minimum is given by

$$\min_{x \in S} f(\boldsymbol{x}) = \min_{\boldsymbol{x}} \max_{\lambda} \max_{\alpha, \alpha \geq 0} L(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\alpha})$$

- We will the proof and details of this when we come to later weeks (SVM).
- We will not be using this during the Deep Learning portions of the course.