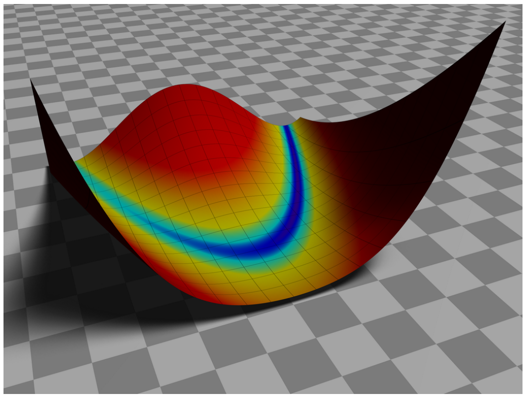
Penalty Functions

Lecture 17



ME EN 575
Andrew Ning
aning@byu.edu

Outline

Penalty Methods

Augmented Lagrangian

Wing Example

Recall the Lagrangian:

$$\mathcal{L}(x, \lambda) = f(x) + \lambda \hat{c}(x)$$

and the KKT conditions:

$$\begin{aligned}
&\frac{\partial f}{\partial x_i} - \sum_{j=1}^{\hat{m}} \hat{\lambda}_j \frac{\partial \hat{c}_j}{\partial x_i} - \sum_{k=1}^{m} \lambda_k \frac{\partial c_k}{\partial x_i} = 0, \quad i = 1, \ldots, n \\
&\hat{c}_j = 0, \quad j = 1, \ldots, \hat{m} \\
&c_k - s_k^2 = 0 \quad k = 1, \ldots, m \\
&\lambda_k s_k = 0, \quad k = 1, \ldots, m \\
&\lambda_k \geq 0, \quad k = 1, \ldots, m
\end{aligned}$$

Penalty Methods

Penalty methods are not often used anymore, but they are the easiest to understand so we start with them.

Basic Idea: translate constrained problem into a sequence of unconstrained problems. We do this by penalizing constraint violations and adding them to the objective. Thus, constraint violations become undesirable in the objective.

Define a new objective:

$$F(x) = f(x) + \mu P(x)$$

Notice the similarity to the Lagrangian, except for we supply μ and its just one value.

Quadratic Penalty

Equality constrained:

$$F(x; \mu_{eq}) = f(x) + \frac{\mu_{eq}}{2} \sum_i \hat{c}_i(x)^2$$

Inequality constrained:

$$F(x; \mu_{in}) = f(x) + \frac{\mu_{in}}{2} \sum_i max[0, c_i(x)]^2$$

As $\mu \rightarrow \infty$ should recover exact solution.

Procedure:

1. Provide starting guess for μ.
2. Solve unconstrained problem.
3. Increase μ.
4. Go back to 1 and repeat until converged.

Problems:

- ill-conditioning
- some formulations are not smooth
- may be unbounded below
- solution is always somewhat infeasible

Example

$$\begin{aligned} \text{minimize} \quad & x_1 + x_2 \\ \text{subject to} \quad & x_1^2 + x_2^2 = 8 \end{aligned}$$

(see Penalty notebook)

Many other penalty methods exist, some of which are nonsmooth (l_1 norm).

They all require managing μ carefully, and dealing with ill-conditioning.

Again, keep in mind better methods now exist and these are not used anymore, but they are still instructive.

Augmented Lagrangian

The quadratic penalty method always produce infeasible results. We can show that each constraint is approximately:

$$c(x) \approx \frac{\lambda^*}{\mu}$$

Thus, we cannot satisfy active constraints without making $\mu \to \infty$.

We can use that information to do better. Let's try to estimate the Lagrange multiplier. We add the quadratic penalty to an estimate of the Lagrangian and call it the augmented Lagrangian:

$$\mathcal{L}(x, \lambda; \mu) = f(x) + \sum_i \lambda_i c_i(x) + \frac{\mu}{2} \sum_i c_i(x)^2$$

Look at optimality conditions of this problem:

$$\nabla_x \mathcal{L}(x, \lambda; \mu) = \nabla f + \sum_i [\lambda_i + \mu c_i] \nabla c_i$$

Compare to actual optimality conditions of the constrained problem:

$$\nabla f + \lambda^* \nabla c$$

Suggests:

$$\lambda^* \approx \lambda_i + \mu c_i(x)$$

Rearrange:

$$c(x) \approx \frac{1}{\mu}(\lambda^* - \lambda_i)$$

Compare to previous result:

$$c(x) \approx \frac{\lambda^*}{\mu}$$

Thus, we can reduce error by increasing μ or by providing an estimate of λ_i that is closer to its true value.

Using:

$$\lambda^* \approx \lambda_i + \mu c_i(x)$$

Suggests an update rule for our estimate:

$$\lambda_i^{k+1} = \lambda_i^k + \mu_k c_i(x_k)$$

Benefits:

- Can assure convergence without increasing $\mu \to \infty$, which addresses issues of ill-conditioning (update μ less frequently).
- Gives us two ways to improve accuracy instead of one.

Wing Example

Look at HW 4.

In Class Demo: Wing example

Highlights some of the challenges you will face in your projects.