数值优化（Numerical Optimization）学习系列-惩罚和增广拉格朗日方法（Augmented Lagrangian Methods）

下一步　于 2015-12-27 18:56:54 发布

概述

求解带约束的最优化问题，一类很重要的方法就是将约束添加到目标函数中，从而转换为一系列子问题进行求解，最终逼近最优解。关键问题是如何将约束进行转换。本节主要介绍
1. 二次惩罚方法
2. 非平滑惩罚方法
3. 增广 拉格朗日 方法

二次惩罚方法

动机

带约束问题如果转换为目标函数加上一个对约束的惩罚项，则问题转换为一个无约束问题。
转换后的问题可以通过惩罚项的系数进行控制，一个比较常见的惩罚函数就是二次惩罚。

等式约束的最优化问题

等式约束问题可以表示为

$$min\ f(x) s.t\ c_i(x) = 0,\ i \in \mathcal{E}$$

添加一个二次惩罚项，则有

$$Q(x;\mu) = f(x) + \frac{\mu}{2}\sum_{i\in\mathcal{E}} c_i^2(x)$$

其中 μ是惩罚参数，直观上只要增加惩罚参数的值就可以逼近原始问题的最优解。
在实际中，对于某个惩罚参数 μ只要几步无约束最优化问题，不需要寻找最优解。

一般化约束最优化问题

一般化约束最优化问题表示为

$$\begin{aligned} & min f(x) \\ & s.t\ c_i(x) = 0\ i \in \mathcal{E} \\ & \quad c_i(x) \geq 0\ i \in \mathcal{I} \end{aligned}$$

添加惩罚项系数结果为

$$Q(x;\mu) = f(x) + \frac{\mu}{2}\sum_{i\in\mathcal{E}} c_i^2(x) + \frac{\mu}{2}\sum_{i\in\mathcal{I}}([c_i(x)]^-)^2$$

其中 $c_i(x)^-$ 表示当该值大于0时，结果为0，否则为 $-c_i(x)$

二次惩罚项通用框架

Framework 17.1 (Quadratic Penalty Method).

Given $\mu_0 > 0$, a nonnegative sequence $\{\tau_k\}$ with $\tau_k \to 0$, and a starting point x_0^s;
for $k = 0, 1, 2, \ldots$
    Find an approximate minimizer x_k of $Q(\cdot;\mu_k)$, starting at x_k^s,
        and terminating when $\|\nabla_x Q(x;\mu_k)\| \leq \tau_k$;
    if final convergence test satisfied
        stop with approximate solution x_k;
    end (if)
    Choose new penalty parameter $\mu_{k+1} > \mu_k$;
    Choose new starting point x_{k+1}^s;
end (for)

> 1.参数μ的选择可以根据无约束问题的优化难度进行确定，如果很容易优化则可以$\mu_{k+1} = \mu_k$，否则可以选择$\mu_{k+1} = \mu_k$
> 2. 定理：如果转换后的问题$Q(x;\mu_k)$每一步都计算最优解，并且当$\mu_k \to \infty$时能够接近原始问题的最优解。

非平滑惩罚函数

有些惩罚函数是精确的，即惩罚项参数μ达到一定值时转换后的问题的最优解就是原始问题的最优解，其中l1惩罚项就是精确的，表示如下

$$\phi_1(x;\mu) = f(x) + \mu\sum_{i\in\mathcal{E}} |c_i(x)| + \mu\sum_{i\in\mathcal{E}} |c_i(x)|^-$$

通用求解框架

Framework 17.2 (Classical ℓ_1 Penalty Method).
Given $\mu_0 > 0$, tolerance $\tau > 0$, starting point x_0^s;
for $k = 0, 1, 2, \ldots$
    Find an approximate minimizer x_k of $\phi_1(x; \mu_k)$, starting at x_k^s;
    if $h(x_k) \leq \tau$
        stop with approximate solution x_k;
    end (if)
    Choose new penalty parameter $\mu_{k+1} > \mu_k$;
    Choose new starting point x_{k+1}^s;
end (for)

增广拉格朗日方法

动机

增广拉格朗日方法在拉格朗日方法的基础上添加了二次惩罚项，从而使得转换后的问题能够更容易求解，不至于因条件数变大不好求。则转换后的问题为

$$\mathcal{L}(x, \lambda; \mu) = f(x) - \sum_{i \in \mathcal{E}} \lambda_i c_i(x) + \frac{\mu}{2} \sum_{i \in \mathcal{E}} c_i(x)^2$$

在第K步迭代过程中，固定惩罚项参数 μ和 λ^k，此时优化x，根据最优化条件有

$$\nabla_x \mathcal{L} = \nabla f(x) - \sum_{i \in \mathcal{E}} (\lambda_i^k - \mu_k c_i(x)) \nabla c_i(x) = 0$$

对比最优性条件，应该有 $\nabla f(x^*) = 0; \lambda^* = \lambda_i^k - \mu_k c_i(x)$，从而很自然的可以将 $\lambda^{k+1} = \lambda_i^k - \mu_k c_i(x)$

等式约束通用框架

Framework 17.3 (Augmented Lagrangian Method-Equality Constraints).

Given $\mu_0 > 0$, tolerance $\tau_0 > 0$, starting points x_0^s and λ^0;

for $k = 0, 1, 2, \ldots$

Find an approximate minimizer x_k of $\mathcal{L}_A(\cdot, \lambda^k; \mu_k)$, starting at x_k^s, and terminating when $\|\nabla_x \mathcal{L}_A(x_k, \lambda^k; \mu_k)\| \leq \tau_k$;

if a convergence test for (17.1) is satisfied

stop with approximate solution x_k;

end (if)

Update Lagrange multipliers using (17.39) to obtain λ^{k+1};

Choose new penalty parameter $\mu_{k+1} \geq \mu_k$;

Set starting point for the next iteration to $x_{k+1}^s = x_k$;

Select tolerance τ_{k+1};

end (for)

实际应用

在实际中，增广拉格朗日方法可以很有效的处理边界约束和线性约束最优化问题。

总结

了解通过将约束转换为惩罚项添加到目标函数上的方法，了解增广拉格朗日方法的动机。