数值优化（Numerical Optimization）学习系列-非线性方程（Nonlinear Equation）

下一步 于 2015-12-27 18:52:46 发布

概述

> 实际中很多应用不是寻找最优解，而是寻找一个根满足给定的约束条件，如果有n个 非线性 等式约束，就是本节介绍的非线性方程问题，本节主要介绍
> 1. 非线性方程的问题形式
> 2. 非线性方程的求解算法
> 3. 总结

非线性方程的问题形式

问题形式，寻找满足n个非线性等式的根，即

$$r(x)=0$$

其中 $r(x)=[r_1(x),...r_n(x)]^T$
该问题可以转换为求解 $min\sum_{i=1...n} r_i^2(x)$，根据最优化条件可知，最优解应该满足： $\nabla f(x^*)=0 => J(x^*)r(x^*)=0$。

> 1.该最小化问题等价于最小二乘问题，可以采用非线性最小二乘的求解算法，不同点在于有n个非线性等式。
> 2. 根据最优性条件，如果雅克比矩阵J(x)是非奇异的，则最优解就是该非线性方程的根；否则最优解可能不会满足r(x)=0
> 3. 如果对于病态的J(x)，即如果秩为n-1或者n-2则最优化算法得到的解会比较差。
> 4. 最优化问题往往最优解只能得到一个，对于非线性方程可能存在唯一解、有限解、无穷解或者无解。

非线性方程的求解算法

该问题求解算法，主要思路是转换为最优化问题，利用类牛顿方法进行求解。

牛顿方法

类似于最优化问题的牛顿方法，寻找牛顿方向，然后不断迭代。算法如下：

Algorithm 11.1 (Newton's Method for Nonlinear Equations).

Choose x_0;

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

Calculate a solution p_k to the Newton equations

$$J(x_k)p_k = -r(x_k);$$

$x_{k+1} \leftarrow x_k + p_k$;

end (for)

> 算法流程
> 1. 每次寻找满足牛顿方程的方向进行迭代。
> 2. 牛顿方程来源，对r(x)进行泰勒展开，有
>
> $$r(x+p) = r(x) + J(x)p$$
>
> 相当于定义一个线性模型近似该目标，即
>
> $$M(p) = r(x) + J(x)p$$
>
> 3. 牛顿方程来源二，看做是最优问题
>
> $$minf(x) = \frac{1}{2}||r(x)||^2$$
>
> 对应的牛顿方程为
>
> $$\nabla^2 f(x)p = -\nabla f(x)$$
>
> ，对应于该问题为
>
> $$(J(x)^T J(x)p + m(x)p + J(x)^T r(x)) = 0$$
>
> ，其中r(x)的二阶梯度一般可以省略（参加最小二乘问题）,即
>
> $$J(x)^T (J(x)p + r(x)) = 0$$
>
> ，如果J(x)非奇异有

> $$J(x)p + r(x) = 0$$
>
> 4. 对比牛顿方程和线性最小二乘问题，可以看出该方程时如下最小二乘问题的优化目标，即 $minf(x) = \frac{1}{2}||r_k + J_k p||^2$，因此可以利用最小二乘相关算法进行求解该等式
> 5. 该算法的收敛速度为超线性收敛。

该算法的局限性

> 1. 当初始点选取不当，离根较远时，算法表现不好。当J奇异时，不一定能找到解。
> 2. 雅克比矩阵不一定容易求解
> 3. 当n比较大时，牛顿方程不好求解。
> 4. 当J是奇异矩阵是，结果不可预知。

改进算法

非精确牛顿方法

不同于牛顿方法，之间寻找满足Jp=-r的牛顿方向，该方法寻找满足如下条件的方向

$$||r_k + J_k P_k|| \leq \eta ||r_k||$$

其中 η是force sequence，即非递增序列。
算法如下

Framework 11.2 (Inexact Newton for Nonlinear Equations).
Given $\eta \in [0, 1)$;
Choose x_0;
for $k = 0, 1, 2, \ldots$
    Choose forcing parameter $\eta_k \in [0, \eta]$;
    Find a vector p_k that satisfies (11.17);
    $x_{k+1} \leftarrow x_k + p_k$;
end (for)

1. 算法的收敛速度依赖与参数η的选取
2. 如果η充分小则为线性收敛；如果$\eta \to 0$则为超线性收敛；如果$\eta_k = o(||r_k||)$二次收敛。

Broyden方法

这是一类拟牛顿方法，思路如下
1. 迭代过程中，近似模型为$M(p) = r(x) + B_k p$，其中B为雅克比矩阵的近似
2. 根据拟牛顿方程，定义如下

$$s_k = x_{k+1} - x_k, y_k = r(x_{k+1}) - r(x_k), y_k = B_{k+1} s_k$$

其中该方程也成为割线方程（secant equation）
3. Broyden方法提供了一种比较实用的更新方法，即

$$B_{k+1} = B_k + \frac{(y_k - B_k s_k) s_k^T}{s_k^T s_k}$$

4. 该方法收敛速度为超线性收敛

算法步骤为

Algorithm 11.3 (Broyden).

Choose x_0 and a nonsingular initial Jacobian approximation B_0;

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

Calculate a solution p_k to the linear equations

$$B_k p_k = -r(x_k);$$

Choose α_k by performing a line search along p_k;

$x_{k+1} \leftarrow x_k + \alpha_k p_k$;

$s_k \leftarrow x_{k+1} - x_k$;

$y_k \leftarrow r(x_{k+1}) - r(x_k)$;

Obtain B_{k+1} from the formula (11.28);

end (for)

Tensor方法

张量方法，相当于在线性近似模型的基础上，添加了表示非线性参数。即$M(p) = r(x) + Jp + Tpp$其中T是一个三维数据相当于r(x)的二阶梯度。

实际算法

在实际应用中，在牛顿算法的基础上，使用线搜索和信赖域进行算法优化。

> 1.目标函数（Merit Func）不仅仅可以选取平方形式，也可以选取L1正则，即$minf(x) = |r(x)|$
> 2. 线搜索，更新步骤为$x_{k+1} = x_k + \alpha p_k$
> 3. 信赖域，优化目标为 $minf(x) = \frac{1}{2}||r_k + J_k p||^2$； $p \leq \Delta_k$
> 4. 在实际中为解决雅克比矩阵奇异或者病态的问题，搜索方向选取为$p_k = -(J_k^T J + \lambda_k I)^{-1} J_k^T r_k$，相当于目标函数添加了L2正则项

连续方法（Continuation Methods）

该方法的思路，如果原始问题求解比较复杂，将其转换为一系列简单的问题，其中一类简单的转换思路为，Homotopy Map：

$$H(x, \lambda) = \lambda r(x) + (1 - \lambda)(x - a)$$

当 $\lambda = 0$时，最优解为$x = 1$，否则当$\lambda = 1$是，$H = 0$等价于$r(x) = 0$。
如果从$0 - 1$逐渐修改λ的值，计算$H(x, \lambda) = 0$的根，最终可以得到$r(x) = 0$

总结

通过该节的学习，可以了解到

1. 非线性方程如何转换为最优化问题
2. 牛顿方法以及变种求解非线性方程