优化理论——二次规划

星空爱好者
优化理论.机器学习

导语

二次规划(QP, Quadratic Programming)定义：目标函数为二次函数，约束条件为线性约束，属于最简单的一种非线性规划。

QP——等式约束

一个标准的等式约束QP模型如下

$$
\begin{aligned}
&min\ \frac{1}{2}x^T Hx + g^T x \\
&\ s.t.\ a_i^T x = b_i,\ i = 1, 2, \ldots, m
\end{aligned}
$$

其中 H 是由二阶导构成的Hessian矩阵， g^T 是由梯度构成的Jacobi矩阵，这里的向量都是指列向量， g^T 表示转置成行向量。

其对应的Lagrange函数为

$$
L(x, \lambda) = \frac{1}{2}x^T Hx + g^T x + \sum_{i=1}^{m} \lambda_i (a_i^T x - b_i)
$$

满足KKT条件，即满足

$$
\begin{cases}
\frac{\partial L}{\partial x} = Hx + g + \sum_{i=1}^{m} \lambda_i a_i = 0 \\
\frac{\partial L}{\partial \lambda_i} = a_i^T x - b_i = 0,\ i = 1, 2, \ldots, m
\end{cases}
$$

将其写成矩阵形式

$$
F(x, \lambda) = \begin{bmatrix} Hx + g + \lambda a \\ a^T x - b \end{bmatrix} = 0
$$

其中

$$
\lambda = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \ldots & \lambda_m \end{bmatrix},\ a^T = \begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_m^T \end{bmatrix},\ b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}
$$

这是一个纯线性方程组，易于求解，KKT方程组的解 (x^*, λ^*) ，即为优化模型的解

QP——不等式约束

一个标准的等式不等式联合约束QP原模型如下

$$
\begin{aligned}
&min\ \frac{1}{2}x^T Hx + g^T x \\
&\ s.t.\ a_i^T x = b_i, i \in E \\
&\qquad h_j^T x \le t_j, j \in I
\end{aligned}
$$

$i \in E$ 表示的是 m 个等式约束集合，$i \in I$ 表示的是 n 个不等式约束集合。对于不等式约束下的QP问题此处介绍两种求解方案

内点法

模型的拉格朗日函数为

$$L(x,\lambda,\mu)=\frac{1}{2}x^THx+g^Tx+\sum_{i=1}^m\lambda_i(a_i^Tx-b_i)+\sum_{j=1}^n\mu_j(h_j^Tx-t_i)$$

内点法在之前的系列中我们已经详细介绍，以原始对偶内点法为例，加入微小量 τ 扰动后KKT条件为

$$\begin{cases}Hx+g+\sum_{i=1}^m\lambda_ia_i+\sum_{j=1}^n\mu_jh_j=0\\ a_i^Tx=b_i,\ i=1,\ldots,m\\ h_j^Tx\le t_j\\ \mu_j(h_jx-t_j)=-\tau_j\\ \mu_j\ge 0,\ j=1,\ldots,n\end{cases}$$

为了更快的检查解是否在约束空间内，我们在不等式方程组引入了松弛变量 s_j 则 $s_j=t_j-h_jx$ ，因为判断 $s_j\ge 0$ 要比判断 $h_jx\le t_j$ 方便的多，上式成为

$$\begin{cases}Hx+g+\sum_{i=1}^m\lambda_ia_i+\sum_{j=1}^n\mu_jh_j=0\\ a_i^Tx=b_i,\ i=1,\ldots,m\\ h_j^Tx+s_j=t_j\\ \mu_js_j=\tau\\ \mu_j,s_j\ge 0,\ j=1,\ldots,n\end{cases}$$

将其写成矩阵形式

$$F(x,s,\lambda,\mu)=\begin{bmatrix}Hx+g+\lambda a+\mu h\\ h^Tx+s-t\\ a^Tx-b\\ \mu s-\tau 1\end{bmatrix}=0$$

其中

$$\lambda=\begin{bmatrix}\lambda_1&0&\ldots&0\\0&\lambda_2&\ldots&0\\\vdots&\vdots&\ldots&\vdots\\0&0&\ldots&\lambda_m\end{bmatrix},\ a^T=\begin{bmatrix}a_1^T\\a_2^T\\\vdots\\a_m^T\end{bmatrix},\ b=\begin{bmatrix}b_1\\b_2\\\vdots\\b_m\end{bmatrix}u=\begin{bmatrix}u_1&0&\ldots&0\\0&u_2&\ldots&0\\\vdots&\vdots&\ldots&\vdots\\0&0&\ldots&u_n\end{bmatrix}$$

$$h^T=\begin{bmatrix}h_1^T\\h_2^T\\\vdots\\h_n^T\end{bmatrix}t=\begin{bmatrix}t_1\\t_2\\\vdots\\t_n\end{bmatrix}s=\begin{bmatrix}s_1\\s_2\\\vdots\\s_n\end{bmatrix}$$

牛顿法解方程组

$$F(x_k,s_k,\lambda_k,\mu_k)+F'(x_k,s_k,\lambda_k,\mu_k)(\Delta x_k,\Delta s_k,\Delta\lambda_k,\Delta\mu_k)=0$$

即

$$\begin{bmatrix} H & 0 & a^T & h^T \\ h^T & I & 0 & 0 \\ a^T & 0 & 0 & 0 \\ 0 & u_k & 0 & s_k^T \end{bmatrix} \begin{bmatrix} \Delta x_k \\ \Delta s_k \\ \Delta \lambda_k \\ \Delta \mu_k \end{bmatrix} = - \begin{bmatrix} Hx_k + g + \lambda_k a + \mu_k h \\ h^T x_k + s_k - t \\ a^T x_k - b \\ \mu_k s_k - \tau_k 1 \end{bmatrix}$$

得到 $(\Delta x_k, \Delta s_k, \Delta \lambda_k, \Delta \mu_k)$ 然后更新变量

$$(x_{k+1}, s_{k+1}, \lambda_{k+1}, \mu_{k+1}) = (x_k, s_k, \lambda_k, \mu_k) + \alpha(\Delta x_k, \Delta s_k, \Delta \lambda_k, \Delta \mu_k)$$

同时更新$\tau_{k+1} = -\sigma \sum_{j=1}^n \mu_{j,k} s_{j,k}$，$\sigma \in [0,1]$进入 $k+1$ 次迭代直到方程组的解，并且满足 $\tau_k \leq \epsilon$

积极集法

积极集法属于图形法在QP问题上的扩展，其通过求解有限个等式约束QP问题来解决一般约束下的QP模型。当不等式约束条件不多时，也是一种高效的求解QP算法。

积极集法首先将QP中所有的不等式约束视为等式约束。把不等式约束直接转成等式约束当然是存在问题的，不等式约束存在有效和无效两种情况，而有效无效很容易通过该不等式对应的拉格朗日乘子进行判断。不等式约束的互补松弛条件告诉我们，不等式对应的拉格朗日乘子应当满足 $\lambda_i \geq 0$。

在第 k 次迭代开始时，我们首先检查当前的迭代点 x_k 是否为当前工作集 W_k （有效不等式约束集合）下的最优点。如果不是，我们就通过求解一个等式约束的 QP 命题来得到一个前进方向 p 。在计算p的时候，只关注W_k中的不等式约束并将其转化为等式约束，而忽略其他不等式约束。令：$d = x - x_k$ ，代入原命题得

$$\begin{aligned} &min\ \frac{1}{2}(d+x_k)^T H(d+x_k) + g^T(d+x_k) \\ &s.t.\ a_i^T(d+x_k) = 0, i \in W_k \end{aligned}$$

上式展开后，令 $g_k = Hx_k + g$，$\rho_k = \frac{1}{2}x_k^T H x_k + x_k^T g$ （常数），在第 k 次迭代需要求解的 QP 子命题为

$$\begin{aligned} &min\ \frac{1}{2}d^T Hd + g_k^T d \\ &s.t.\ a_i^T d = 0, i \in W_k \end{aligned}$$

更新：$x_{k+1} = x_k + \beta_k d_k$

添加有效约束：

迭代需要满足 $a_i^T(x_k + \beta_k d_k) \leq b_i$ ，因此步长 β_k 为

$$\beta_k = min\{1, min_{i \notin W_k, a_i^T d_k > 0} \frac{b_i - a_i^T x_k}{a_i^T d_k}\}$$

$\beta_k = 1$：满足所有等式不等式约束，不需要更新工作集，进一步迭代；$\beta_k = 0$：在当前迭代点处有其他的有效约束没有被添加到工作集W_k中；$\beta_k < 1$：也就是说下降方向 p_k 被某条不在工作集 W_k 内的约束阻拦住了，需要将这条约束添加到工作集来构造新的工作集 W_{k+1} 。

删除无效约束：

计算拉格朗日乘子

$$\begin{cases} Hx^* + g = -\sum_{i=1}^{m} a_i \lambda_i \\ \lambda_i (a_i^T x - b_i) = 0 \\ \lambda_i \geq 0,\ i \in W \end{cases}$$

通过上式计算出不等式约束对应的拉格朗日乘子，如果有一个或者多个 λ_i 的值小于0。那么就表明通过去掉工作集的某一条或几条约束，目标函数值可以进一步下降。因此我们会从对应的λ_i值小于 0 的约束中选择一条，将其从工作集W_k中剔除从而构造出新的工作集W_{k+1} 。如果有多于一条的可选约束，那么不同的剔除方法会遵循不同原则，默认去除对应 λ_i 值最小的那条约束。

编辑于 2023-06-30 06:11・IP 属地未知

二次函数 拉格朗日乘子 不等式

5 条评论 默认 最新

郑秋
突然发现看不懂数学公式，尬住
那我岂不是得先去看数学
2022-01-25 回复 1

woooli
一开始写的就有问题。gT一会说是雅可比矩阵，一会说是转置向量
09-11 回复 喜欢

易易紫
请问为什么要微小扰动啊?
07-17 回复 喜欢

boomer
博主你好，内点法整理后的矩阵形式看起来是线性的，能否直接使用线性规划方法求解?
2022-08-09 回复 喜欢

boomer
看漏了哈，公式是非线性的
2022-08-09 回复 喜欢

文章被以下专栏收录

现代优化理论
现代优化理论

轨迹优化代码讲解

均是期望曲线的l（s）,一阶导、二

Manta 发表于玩转Apo...

c_i(x) 的某种近似迭代法，尤其是利用约束函数 c_i(x) 的线性近似。基

稷殿下

的原因 output： 有关优化过程的信

wjr 发表于Matla...

设计呢？ 概念只是看你这个思路。

大大挂