Lecture 8：ALM与ADMM

1. 增广拉格朗日算法 (ALM)

2. 交替方向乘子法(ADMM)

ALM

等式约束优化的乘子罚

- 乘子罚函数【Hestenes,Powell,1969】

$$\min f(x) \quad (1)$$
$$\text{s.t.} \quad c_i(x) = 0, i \in \mathcal{E}$$

$$\Longleftrightarrow$$

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

$$\Big\Downarrow \text{平方惩罚}$$

$$L_\rho(x, y) := f(x) - \sum y_i c_i(x) + \frac{\rho}{2} \sum c_i^2(x)$$

称函数$L_\rho(x, y)$为优化问题的乘子罚函数，也称为增广Lagrange函数.

ALM

精确罚理论

定理1 对优化问题(1),设$\boldsymbol{x}^*$是局部最优解且**二阶充分最优性条件**

$$d^T \nabla_{xx}^2 L(x^*, y^*) d > 0,\ \ \forall d \in \{d \neq 0 : \nabla c_i(x^*)^T d = 0, i \in \mathcal{E}\}$$

成立,其中$\boldsymbol{y}^*$是相应最优Lagrange乘子,设向量组$\{\boldsymbol{\nabla c_i}(\boldsymbol{x}^*), \boldsymbol{i} \in \mathcal{E}\}$线性无关(LICQ),则存在**有限常数**$\rho^* > 0$,使得对$\forall \rho \geq \rho^*$, $\boldsymbol{x}^*$是$\min\limits_{x} \boldsymbol{L}_{\boldsymbol{\rho}}(\boldsymbol{x}, \boldsymbol{y}^*)$的严格局部最优解;特别地,若目标函数$f$为凸函数,所有约束函数为线性的,则对任意$\rho > 0$,问题(1)的最优解都是罚问题$\min\limits_{x} \boldsymbol{L}_{\boldsymbol{\rho}}(\boldsymbol{x}, \boldsymbol{y}^*)$的最优解.另一方面若$\boldsymbol{x}^*$是$\min\limits_{x} \boldsymbol{L}_{\boldsymbol{\rho}}(\boldsymbol{x}, \boldsymbol{y}^*)$的局部最优解且$c_i(x^*) = 0, i \in \mathcal{E}$,则$\boldsymbol{x}^*$是(1)的局部最优解.

补充:二阶最优性条件

二阶最优性条件

约束优化模型

$$
\begin{aligned}
\min \quad & f(x) \\
s.t. \quad & g_i(x) \geq 0, \quad i = 1,2,\cdots,m \text{----不等式约束} \\
& h_j(x) = 0, \quad j = 1,2,\cdots,l. \text{----等式约束} \\
& x \in R^n
\end{aligned}
$$

可行域（集）： $S = \left\{ x \in R^n \middle| \begin{array}{ll} g_i(x) \geq 0, & i = 1,2,\cdots,m \\ h_j(x) = 0, & j = 1,2,\cdots,l \end{array} \right\}$

二阶最优性条件

KKT条件的非充分性

例1

$$\min\ f(x)=x_2$$
$$s.t.\quad g(x)=x_1^2+x_2\geq 0$$

求KKT点，并说明是否为局部最优解。

解：KKT点应满足方程组

$$\begin{cases}\begin{pmatrix}0\\1\end{pmatrix}-w\begin{pmatrix}2x_1\\1\end{pmatrix}=0\\ w(x_1^2+x_2)=0\\ x_1^2+x_2\geq 0\\ w\geq 0\end{cases}\Longrightarrow\quad \begin{aligned}&w=1\\&x_1=x_2=0\end{aligned}$$

$\bar{x}=(0,0)^T$ 是KKT点，但不是局部最优解.

二阶最优性条件

定理14（二阶必要条件）： 设f，$g_i(i=1,\cdots,m)$和$h_j(j=1,\cdots,l)$是二次连续可微函数，$\bar{x}$为一个局部最优解且在$\bar{x}$处LICQ成立，令$\bar{w},\bar{v}$为Lagrange乘子，则

$$\underline{d^T\nabla_x^2 L(\bar{x},\bar{w},\bar{v})d \ge 0, \forall d \in C(\bar{x},\bar{w},\bar{v})} := \left\{ d \in R^n \;\middle|\; \begin{array}{l} \nabla g_i(\bar{x})^T d = 0, i \in I \text{且} \bar{w}_i > 0 \\ \nabla g_i(\bar{x})^T d \ge 0, i \in I \text{且} \bar{w}_i = 0 \\ \nabla h_j(\bar{x})^T d = 0, j = 1,2,\cdots,l \end{array} \right\}.$$

$\nabla_x^2 L(\bar{x},\bar{w},\bar{v})$在临界锥上是半正定的

- 临界锥（**Critical Cone**）：$C(\bar{x},\bar{w},\bar{v})$

$$\forall d \in C(\bar{x},\bar{w},\bar{v}), \text{有} \bar{w}_i\nabla g_i(\bar{x})^T d = 0, \bar{v}_j \nabla h_j(\bar{x})^T d = 0, \forall i, j$$

$\bar{x}$是局部最优解且LICQ成立

$(\bar{x},\bar{w},\bar{v})$是KKT系统的解

$$0 = \nabla_x L(\bar{x},\bar{w},\bar{v}) = \nabla f(\bar{x}) - \sum_{i=1}^m \bar{w}_i \nabla g_i(\bar{x}) - \sum_{j=1}^l \bar{v}_i \nabla h_j(\bar{x})$$

$$d^T\nabla f(\bar{x}) = \sum_{i=1}^m \bar{w}_i d^T \nabla g_i(\bar{x}) + \sum_{j=1}^l \bar{v}_i d^T \nabla h_j(\bar{x}) = 0$$

目标函数值非增非减，临界状态

二阶最优性条件

例2 $\min\ x_1$

$$s.t.\quad x_2 \ge 0$$

$$1-(x_1-1)^2 - x_2^2 \ge 0$$

解：该问题的最优解为$\bar{x}=(0,\ 0)^T$，积极约束集为I=$\{1,2\}$.

* Lagrange函数为：$L(x,w)=x_1-w_1x_2-w_2\left(1-(x_1-1)^2-x_2^2\right)$，其中$w_1$与$w_2$非负

* 由KKT条件$\nabla L(\bar{x},\bar{w})$=0，可知$\begin{cases}1-2\bar{w}_2=0\\ -\bar{w}_1=0\end{cases}$,从而$\bar{w}=(0,1/2)^T$.

* 在$\bar{x}$处LICQ成立： 因为$\nabla g_1(\bar{x})=(0,1)^T$,$\nabla g_2(\bar{x})=(2,0)^T$ 线性无关

* 临界锥为：$C(\bar{x},\bar{w})=\left\{(0,d_2)^T \mid d_2 \ge 0\right\}$

* 二阶必要性条件在$\bar{x}$处成立：因为$\nabla_x^2 L(\bar{x},\bar{w})=\begin{bmatrix}2\bar{w}_2 & 0\\ 0 & 2\bar{w}_2\end{bmatrix}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$正定

二阶最优性条件

二阶充分条件

定理15 若优化问题目标函数与约束函数均二阶连续可微,设$(\bar{\boldsymbol{x}}, \bar{\boldsymbol{w}}, \bar{\boldsymbol{v}})$是KKT点对且满足如下二阶条件

$$d^T \nabla_{xx}^2 L(\bar{\boldsymbol{x}}, \bar{\boldsymbol{w}}, \bar{\boldsymbol{v}}) d > 0, \ \forall d \in \mathrm{C}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{w}}, \bar{\boldsymbol{v}}) \setminus \{0\},$$

则$\bar{\boldsymbol{x}}$是约束优化问题的严格局部最优解.

✓ **用于判断KKT点是否为局部最优解!**

二阶最优性条件

证明： 假设 $\bar{x}$ 不是问题的严格局部极小点, 则存在序列 $\{x^{(k)}\} \subset S$

使得 $x^{(k)} \neq \bar{x}, x^{(k)} \to \bar{x}$ 且 $f(x^{(k)}) \leq f(\bar{x})$, 即

$$f(x^{(k)}) - f(\bar{x}) = \nabla f(\bar{x})^T (x^{(k)} - \bar{x}) + o(\| x^{(k)} - \bar{x} \|) \leq 0.$$

记 $d^{(k)} = \dfrac{x^{(k)} - \bar{x}}{\| x^{(k)} - \bar{x} \|}$, 显然, $\forall k \geq 1, \quad \| d^{(k)} \| \equiv 1.$

根据有界序列的性质可知, 存在一个收敛子列 $\{d^{(k_i)}\} \to d \neq 0$

不妨假设 $d^{(k)} \to d$

因为 $\nabla f(\bar{x})^T d^k + o(1) \leq 0$ 令 $k \to \infty$ 于是有 $d^T \nabla f(\bar{x}) \leq 0$

二阶最优性条件

以下证明 $\nabla f(\bar{x})^T d = 0$

将$g_i(x)$在点$\bar{x}$展开，再令$x = x^{(k)}$,得到

$$g_i(x^{(k)}) = g_i(\bar{x}) + \nabla g_i(\bar{x})^T (x^{(k)} - \bar{x}) + o(\| x^{(k)} - \bar{x} \|)$$

$$\Rightarrow \quad \nabla g_i(\bar{x})^T (x^{(k)} - \bar{x}) + o(\| x^{(k)} - \bar{x} \|) \ge 0 \ (i \in I)$$ 取极限

$$\Rightarrow \quad \nabla g_i(\bar{x})^T d \ge 0 \quad (i \in I)$$

同理可证 $\nabla h_j(\bar{x})^T d = 0, \quad j = 1, 2, \cdots, l$

利用KKT条件, 得到 $\nabla f(\bar{x})^T d = \sum_{i \in I} \bar{w}_i \nabla g_i(\bar{x})^T d + \sum_{j=1}^{l} \bar{v}_j \nabla h(\bar{x})^T d \ge 0,$

因此, $\nabla f(\bar{x})^T d = 0.$

由于 $\nabla f(\bar{x})^T d = \sum_{i \in I} \bar{w}_i \nabla g_i(\bar{x})^T d = 0$

$\therefore \nabla g_i(\bar{x})^T d = 0,$ 若$i \in I$且$\bar{w}_i > 0$ $\quad \Rightarrow \quad d \in C(\bar{x}, \bar{w}, \bar{v})$

二阶最优性条件

$\because x^{(k)} \in S$ 由**Taylor**展开公式, 有

$$L(\bar{x}, \bar{w}, \bar{v}) = f(\bar{x}) \geq f(x^{(k)})$$

$$\geq f(x^{(k)}) - \sum_{i \in I} \bar{w}_i g_i(x^{(k)})^T - \sum_{j=1}^{l} \bar{v}_j h_j(x^{(k)})^T = L(x^{(k)}, \bar{w}, \bar{v})$$

$$\text{而} L(x^{(k)}, \bar{w}, \bar{v}) = L(\bar{x}, \bar{w}, \bar{v}) + \underline{\nabla_x L(\bar{x}, \bar{w}, \bar{v})}^T (x^{(k)} - \bar{x}) \qquad \equiv 0$$

$$+ \frac{1}{2}(x^{(k)} - \bar{x})^T \nabla_x^2 L(\bar{x}, \bar{w}, \bar{v})(x^{(k)} - \bar{x}) + o(\| (x^{(k)} - \bar{x}) \|^2).$$

$$\therefore \frac{1}{2}(x^{(k)} - \bar{x})^T \nabla_x^2 L(\bar{x}, \bar{w}, \bar{v})(x^{(k)} - \bar{x}) + o(\| (x^{(k)} - \bar{x}) \|^2) \leq 0.$$

$\therefore \quad d^T \nabla_x^2 L(\bar{x}, \bar{w}, \bar{v}) d \leq 0.$ 矛盾。

二阶最优性条件

例3 $\min\ f(x)=x_1x_2$

$$s.t.\quad c(x)=x_1^2+x_2^2-1=0$$

解：该问题的Lagrange函数为$L(x,v)=x_1x_2-v\left(x_1^2+x_2^2-1\right)$.

由一阶必要条件$\begin{cases}x_2-2vx_1=0\\ x_1-2vx_2=0\\ x_1^2+x_2^2-1=0\end{cases}$

可知

$$v^*=\frac{1}{2},\quad x^*=\left(\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2}\right)^T \text{ or } x^*=\left(-\frac{\sqrt{2}}{2},-\frac{\sqrt{2}}{2}\right)^T$$

或

$$v^*=-\frac{1}{2},\quad x^*=\left(-\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2}\right)^T \text{ or } x^*=\left(\frac{\sqrt{2}}{2},-\frac{\sqrt{2}}{2}\right)^T$$

当$v^*=\frac{1}{2}$时，对任意$d\neq 0$满足$d_1=-d_2$,有

$$d^T\nabla_x^2L(x^*,v^*)d=-4d_1^2<0$$

由二阶充分条件可知,$\left(\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2}\right)^T$与$\left(-\frac{\sqrt{2}}{2},-\frac{\sqrt{2}}{2}\right)^T$是局部极大值点；

当$v^*=-\frac{1}{2}$时，对任意$d\neq 0$满足$d_1=d_2$,有

$$d^T\nabla_x^2L(x^*,v^*)d=4d_1^2>0$$

由二阶充分条件可知,$\left(-\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2}\right)^T$与$\left(\frac{\sqrt{2}}{2},-\frac{\sqrt{2}}{2}\right)^T$是局部最优解.

二阶最优性条件

无约束优化二阶最优性条件

定理16 若$\bar{\boldsymbol{x}} \in \mathbf{argmin}_{\boldsymbol{x}}\, \boldsymbol{f}(\boldsymbol{x})$, 则$\nabla \boldsymbol{f}(\bar{\boldsymbol{x}}) = \mathbf{0}, \nabla^2 \boldsymbol{f}(\bar{\boldsymbol{x}}) \succcurlyeq \boldsymbol{O}$;

若$\nabla \boldsymbol{f}(\bar{\boldsymbol{x}}) = \mathbf{0}, \nabla^2 \boldsymbol{f}(\bar{\boldsymbol{x}}) \succ \boldsymbol{O}$, 则$\bar{\boldsymbol{x}}$是严格局部最优解.

ALM

精确罚理论

定理1 对优化问题(1),设$\boldsymbol{x}^*$是局部最优解且**二阶充分最优性条件**

$$d^T \nabla_{xx}^2 L(x^*, y^*) d > 0,\ \forall d \in \{d \neq 0 : \nabla c_i(x^*)^T d = 0, i \in \mathcal{E}\}$$

成立,其中$\boldsymbol{y}^*$是相应最优Lagrange乘子,设向量组$\{\boldsymbol{\nabla c_i}(\boldsymbol{x}^*), \boldsymbol{i} \in \mathcal{E}\}$线性无关(LICQ),则存在**有限常数**$\rho^* > 0$,使得对$\forall \rho \geq \rho^*$, $\boldsymbol{x}^*$是$\min\limits_{x} \boldsymbol{L_\rho}(\boldsymbol{x}, \boldsymbol{y}^*)$的严格局部最优解;特别地,若目标函数$f$为凸函数,所有约束函数为线性的,则对任意$\rho > 0$,问题(1)的最优解都是罚问题$\min\limits_{x} \boldsymbol{L_\rho}(\boldsymbol{x}, \boldsymbol{y}^*)$的最优解.另一方面若$\boldsymbol{x}^*$是$\min\limits_{x} \boldsymbol{L_\rho}(\boldsymbol{x}, \boldsymbol{y}^*)$的局部最优解且$c_i(x^*) = 0, i \in \mathcal{E}$, 则$\boldsymbol{x}^*$是 (1)的局部最优解.

证明

ALM

ALM的更新准则

$$x^{k+1} \in \operatorname{argmin}_x L_{\rho_k}(x, y^k)$$

$$\Downarrow$$ 【一阶条件】

$$\nabla L_{\rho_k}\left(x^{k+1}, y^k\right) = \nabla f\left(x^{k+1}\right) - \sum_{i \in \varepsilon} \left[y_i^k - \rho_k c_i\left(x^{k+1}\right)\right] \nabla c_i\left(x^{k+1}\right) = 0$$

$$\Downarrow$$

乘子校正公式

$$\boxed{y_i^{k+1} = y_i^k - \rho_k c_i\left(x^{k+1}\right)}$$

$$\nabla f(x^*) - \sum_{i \in \varepsilon} y_i^* \nabla c_i(x^*) = 0$$ 【KKT条件】

Q：可行性?

ALM

ALM的更新准则

算法停机准则： $\|c(x^{k+1})\|_\infty \le \epsilon$

要使得$c_i(x^{k+1}) \to 0$,可以让乘子$y^k \to y^*$,也可以让罚因子不断增加，不一定需要$\rho_k \to +\infty$.

⇧

乘子校正公式

$$\boxed{y_i^{k+1} = y_i^k - \rho_k c_i(x^{k+1})} \implies c_i(x^{k+1}) \approx -(y^* - y^k)/\rho_k$$

ALM

ALM（乘子罚函数法）

迭代框架

步1 取$\rho_0 > 0, y_0 = 0, \gamma > 1$，取初始点$x_0$, $\epsilon > 0$, $k = 1$

步2 以x^{k-1}为初始点求罚函数$L_{\rho_{k-1}}(x, y^{k-1})$的极小值点$x^k$

步3 若$\|c(x^k)\|_\infty \leq \epsilon$,算法终止, 否则转到下一步

步4 若$\|c(x^k)\|_\infty \geq \|c(x^{k-1})\|_\infty$,令$\rho_k = \gamma\rho_{k-1}$, $y^k = y^{k-1}$,令$k = k + 1$
转步2

步5 若$\rho_k > \rho_{k-1}$,或 $\|c(x^k)\|_\infty \leq \frac{1}{4}\|c(x^{k-1})\|_\infty$,令$\rho_k = \rho_{k-1}$, 更新
Lagrange乘子$y_i^{k+1} = y_i^k - \rho_k c_i(x^k)$, $i \in \mathcal{E}$,令$k = k + 1$, 转步2;
否则令$\rho_k = \gamma\rho_{k-1}$, $y^k = y^{k-1}$,令$k = k + 1$

ALM与PPA

复合优化问题：

$$\min_x g(x) + h(y) \quad \text{s.t.} \quad Ax - y = 0 \qquad \text{(P)}$$

其中，g, h为适当闭凸函数，$A \in \mathbb{R}^{m\times n}$

$$-A^T u^{k+1} \in \partial g\left(x^{k+1}\right)$$

$$\Updownarrow$$

ALM for (P):

$$\boxed{\left(x^{k+1}, y^{k+1}\right) = \operatorname{argmin}_{x,y} g(x) + h(y) - \langle u^k, y - Ax\rangle + \frac{\rho}{2}\|Ax - y\|^2}$$

$$\Updownarrow$$

$$u^{k+1} = u^k + \rho\left(Ax^{k+1} - y^{k+1}\right)$$

$$u^{k+1} \in \partial h\left(x^{k+1}\right)$$

ALM与PPA

对偶问题(Fenchel Conjugate Dual)：

$$\min_u \varphi(u) := g^*\left(-A^T u\right) + h^*(u) \qquad \text{(D)}$$

其中 φ 为适当闭凸函数

其他转化方式?

最优性分析：

$$\begin{aligned} 0 \in \partial\varphi(u) &\Leftrightarrow 0 \in \rho\partial\varphi(u) \\ &\Leftrightarrow u \in u + \rho\partial\varphi(u) = (I + \rho\partial\varphi)(u) \\ &\Leftrightarrow u = (I + \rho\partial\varphi)^{-1}(u) = \mathbf{prox}_{\rho\varphi}(u) \end{aligned}$$

不动点迭代：

$$u^{k+1} = \mathrm{Prox}_{\rho\varphi}\left(u^k\right)$$

(Proximal Point Algorithm, PPA)

ALM与PPA

Primal ALM $\Longleftrightarrow$ Dual PPA

对偶问题： $\min_u \varphi(u) := g^*(-A^T u) + h^*(u)$ (D)

PPA for (D): $u^{k+1} = \mathrm{Prox}_{\rho\varphi}(u^k)$

$$\Updownarrow$$

$$u^k - u^{k+1} \in -\rho\left(\boxed{\underbrace{A\partial g^*(-A^T u^{k+1})}_{x^{k+1}}} - \underbrace{\partial h^*(u^{k+1})}_{y^{k+1}}\right)$$

$$\Updownarrow$$

$$u^{k+1} = u^k + \rho\left(Ax^{k+1} - y^{k+1}\right)$$

$$u^{k+1} \in \partial h(x^{k+1})$$

ALM for (P):

$$-A^T u^{k+1} \in \partial g(x^{k+1})$$

ALM与PPA

关键性质： 若g为适当闭凸函数, 则$x \in \partial g^*(y) \iff y \in \partial g(x)$

证明：$x \in \partial g^*(y) \iff y \in \mathrm{argmax}_u \{\langle u, x\rangle - g^*(u)\}$

$$\iff g(x) = g^{**}(x) = \langle y, x\rangle - g^*(y)$$

$$\iff g^*(y) = \langle y, x\rangle - g(x)$$

$$\iff y \in \partial g(x)$$

- 凸优化一阶最优性条件
- $g(x) = g^{**}(x)$

ALM

不等式约束优化的乘子罚

■ 乘子罚函数【Rockafellar,1973】

$$\min f(x) \quad \text{s.t.} \quad c_i(x) \geq 0, i \in I \qquad (2) \qquad \Longleftrightarrow \qquad \min f(x) \quad \text{s.t.} \quad c_i(x) - s_i = 0, s_i \geq 0, i \in \mathcal{E}$$

$\Downarrow$ 乘子罚

$$\min_{s\geq 0,x} f(x) - \sum y_i(c_i(x) - s_i) + \frac{\rho}{2}\sum(c_i(x) - s_i)^2$$

$$= \min_x \min_{s\geq 0} f(x) - \sum y_i(c_i(x) - s_i) + \frac{\rho}{2}\sum(c_i(x) - s_i)^2$$

$\Downarrow$ 最优解：$s_i = \max\{0, c_i(x) - \frac{y_i}{\rho}\}$

$$= \min_x L_\rho(x, y) := f(x) + \sum \psi(c_i(x), y_i, \rho)$$

ALM

不等式约束优化的乘子罚

■ 乘子罚函数【Rockafellar,1973】

$$\psi(t,\sigma,\rho)=\begin{cases}-\sigma t+\dfrac{\rho}{2}t^2, & \text{若} t\le\dfrac{\sigma}{\rho}\\ -\dfrac{\sigma^2}{2\rho}, & \text{其他}\end{cases}$$

称函数$L_\rho(x,y)=f(x)+\sum\psi(c_i(x),y_i,\rho)$为不等式约束优化问题的乘子罚函数，或增广Lagrange函数.

ALM

ALM更新规则

$$\nabla L_{\rho_k}\left(x^{k+1}, y^k\right) = \nabla f\left(x^{k+1}\right) - \sum_{i \in I, y_i^k - \rho_k c_i\left(x^k\right) \geq 0} \left[y_i^k - \rho_k c_i\left(x^{k+1}\right)\right] \nabla c_i\left(x^{k+1}\right) = 0$$

$$\Downarrow$$

$$\nabla f\left(x^{k+1}\right) - \sum_{i \in I} \left[y_i^k - \rho_k c_i\left(x^{k+1}\right)\right]_+ \nabla c_i\left(x^{k+1}\right) = 0$$

$$\Downarrow$$

$$\nabla f\left(x^*\right) - \sum_{i \in \varepsilon} y_i^* \nabla c_i\left(x^*\right) = 0$$

乘子校正公式

$$y_i^{k+1} = \left[y_i^k - \rho_k c_i\left(x^{k+1}\right)\right]_+$$

算法终止准则： $\left\|c_i\left(x^k\right) - s^k\right\|_\infty \leq \epsilon$

ALM

ALM的特点

- 初始点的选取：任意选取
- 迭代过程中的任意点一般是不可行的
- 当迭代点靠近最优解时，随着罚因子的不断增加，子问题的稳定性较好（在一定条件下具有精确罚性质）
- 子问题的精确求解难度

ADMM

分块结构的约束优化问题：

$$\begin{aligned} &\min f_1(x_1) + f_2(x_2) \\ &\text{s.t.} \quad A_1 x_1 + A_2 x_2 = b \end{aligned} \qquad (3)$$

其中，f_1, f_2为适当闭凸函数, $A_1 \in \mathbb{R}^{p\times n}, A_2 \in \mathbb{R}^{p\times m}, b \in \mathbb{R}^p$

问题结构：目标函数可分离成两块, 但变量被线性约束结合在一起

ADMM

例1 复合优化问题

$$\min_x f_1(x) + f_2(x) \quad \Longleftrightarrow \quad \begin{aligned} &\min_{x,z} f_1(x) + f_2(z) \\ &\text{s.t.} \quad x - z = 0 \end{aligned}$$

例2 凸集约束优化问题

$$\begin{aligned} &\min_x f(x) \\ &\text{s.t.} \quad Ax \in C \end{aligned} \quad \Longleftrightarrow \quad \begin{aligned} &\min_{x,z} f(x) + \mathrm{I}_C(z) \\ &\text{s.t.} \quad Ax - z = 0 \end{aligned}$$

ADMM

例3 全局一致性问题（global consensus problem)

$$\min_x \sum_{i=1}^{N} f_i(x) \quad \Longleftrightarrow \quad \min_{x_i,z} \sum_{i=1}^{N} f_i(x_i)$$

$$\text{s.t.} \quad x_i - z = 0, i = 1, \dots, N$$

例4 共享问题(Sharing problem)

$$\min_{x_i} \sum_{i=1}^{N} f_i(x_i) + g\left(\sum_{i=1}^{N} x_i\right) \quad \Longleftrightarrow \quad \min_{x_i,z} \sum_{i=1}^{N} f_i(x_i) + g\left(\sum_{i=1}^{N} z_i\right)$$

$$\text{s.t.} \quad x_i - z_i = 0, i = 1, \dots, N$$

ADMM

分块结构的约束优化问题：

$$\begin{aligned} &\min f_1(x_1) + f_2(x_2) \\ &\text{s.t.}\ \ A_1x_1 + A_2x_2 = b \end{aligned} \qquad (3)$$

其中：f_1, f_2为适当闭凸函数, $A_1 \in \mathbb{R}^{p\times n}, A_2 \in \mathbb{R}^{p\times m}, b \in \mathbb{R}^p$

- **增广Lagrange函数：**

$$L_\rho(x_1, x_2, y) := f_1(x_1) + f_2(x_2) + y^T(A_1x_1 + A_2x_2 - b)$$

$$+\frac{\rho}{2}\|A_1x_1 + A_2x_2 - b\|_2^2$$

ADMM

ALM 迭代

$$\begin{cases} \left(x_1^{k+1}, x_2^{k+1}\right) = \operatorname{argmin}_{x_1,x_2} \ L_{\rho_k}\left(x_1, x_2, y^k\right) \\ y^k = y^k - \tau\rho_k\left(A_1x_1^{k+1} + A_2x_2^{k+1} - b\right) \end{cases}$$

$\tau \in (0, 2)$

$\Downarrow$ 交替极小

ADMM 迭代

$$\begin{cases} x_1^{k+1} = \operatorname{argmin}_{x_1} \ L_{\rho_k}\left(x_1, x_2^k, y^k\right) \\ x_2^{k+1} = \operatorname{argmin}_{x_2} \ L_{\rho_k}\left(x_1^{k+1}, x_2, y^k\right) \\ y^k = y^k - \tau\rho_k\left(A_1x_1^{k+1} + A_2x_2^{k+1} - b\right) \end{cases}$$

$\tau \in \left(0, \frac{1+\sqrt{5}}{2}\right)$

ADMM

分块结构的约束优化问题：

$$
\begin{aligned}
&\min f_1(x_1) + f_2(x_2) \\
&\text{s.t.} \quad A_1 x_1 + A_2 x_2 = b
\end{aligned}
\tag{3}
$$

其中：f_1, f_2为适当闭凸函数，$A_1 \in \mathbb{R}^{p\times n}, A_2 \in \mathbb{R}^{p\times m}, b \in \mathbb{R}^p$

- **KKT 条件**
$$
\begin{cases}
0 \in \partial L(x_1^*, x_2^*, y^*) = \begin{pmatrix} \partial f_1(x_1^*) + A_1^T y^* \\ \partial f_2(x_2^*) + A_2^T y^* \end{pmatrix} \\
A_1 x_1^* + A_2 x_2^* = b
\end{cases}
\tag{4}
$$

其中：$L(x_1, x_2, y) := f_1(x_1) + f_2(x_2) + y^T(A_1 x_1 + A_2 x_2 - b)$为Lagrange 函数.

ADMM

最优性条件

定理2 设凸优化问题(3)满足Slater条件

$$\exists\left(\boldsymbol{x}_1^0, \boldsymbol{x}_2^0\right) \in \textbf{int dom } \boldsymbol{f}_1 \times \textbf{int dom } \boldsymbol{f}_2 \quad \text{s.t. } \boldsymbol{A}_1\boldsymbol{x}_1^0 + \boldsymbol{A}_2\boldsymbol{x}_2^0 = \boldsymbol{b}$$

成立. 则任意最优解$\boldsymbol{x}^*$均为(3)的KKT点,即存在$\boldsymbol{y}^* \in \mathbb{R}^{\boldsymbol{p}}$使得$(\boldsymbol{x}^*, \boldsymbol{y}^*)$满足KKT条件(4).

- KKT点与最优解等价 (在Slater条件下)
- KKT条件可用于设置算法停机准则与分析收敛性

ADMM

最优性分析

$$\min f_1(x_1) + f_2(x_2)$$
$$\text{s.t.}\ \ A_1x_1 + A_2x_2 = b$$

(3)

$$\begin{cases} \begin{pmatrix} 0 \\ 0 \end{pmatrix} \in \begin{pmatrix} \partial f_1(x_1^*) + A_1^T y^* \\ \partial f_2(x_2^*) + A_2^T y^* \end{pmatrix} \\ A_1x_1^* + A_2x_2^* = b \end{cases}$$

$$\Uparrow$$

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} \in \begin{pmatrix} \partial f_1(x_1^*) \\ \partial f_2(x_2^*) \end{pmatrix} + \boxed{\partial \mathrm{I}_\Omega(x_1^*, x_2^*)}$$

$$\Uparrow$$ Slater

$$\Downarrow$$

$$\min f(x_1, x_2) := f_1(x_1) + f_2(x_2) + \mathrm{I}_\Omega(x_1, x_2) \Longrightarrow (0, 0) \in \partial f(x_1^*, x_2^*)$$

$$\Omega := \{(x_1, x_2): A_1x_1 + A_2x_2 = b\}$$

ADMM

对偶问题

$$\sup_{y\in\mathbb{R}^p}\left\{\vartheta(y) := \inf_{x_1,x_2} L(x_1,x_2,y)\right\}$$

$$= -\inf_{y\in\mathbb{R}^p} \underbrace{b^T y + f_1^*\left(-A_1^T y\right)}_{f(y)} + \underbrace{f_2^*\left(-A_2^T y\right)}_{h(y)} \qquad (5)$$

$$\inf_{x_1,x_2} L(x_1,x_2,y)$$
$$= \inf_{x_1}\{f_1(x_1) + y^T A_1 x_1\} + \inf_{x_2}\{f_2(x_2) + y^T A_2 x_2\} - b^T y$$
$$= -\sup_{x_1}\{\langle -A_1^T y, x_1\rangle - f_1(x_1)\} - \sup_{x_2}\{\langle -A_2^T y, x_2\rangle - f_2(x_2)\} - b^T y$$
$$= -f_1^*\left(-A_1^T y\right) - f_2^*\left(-A_2^T y\right) - b^T y$$

ADMM

最优性条件

定理3 设$(\boldsymbol{x}^*, \boldsymbol{y}^*)$满足KKT条件(4), 则$\boldsymbol{x}^*$为原问题(3)的最优解,$\boldsymbol{y}^*$为对偶问题(5)的最优解.

分析

$$\mathbf{0} \in T(x_1^*, x_2^*, y^*)$$

$$\Updownarrow$$

KKT 系统（4）

$$\boxed{\begin{cases} -A_1^T y^* \in \partial f_1(x_1^*) \\ -A_2^T y^* \in \partial f_2(x_2^*) \\ A_1 x_1^* + A_2 x_2^* = b \end{cases}}$$

$$-A_1^T y^* \in \partial f_1(x_1^*) \Longleftrightarrow x_1^* \in \partial f_1^*(-A_1^T y^*)$$

$$-A_2^T y^* \in \partial f_2(x_2^*) \Longleftrightarrow x_2^* \in \partial f_2^*(-A_1^T y^*)$$

$$\Rightarrow \mathbf{0} \in \partial f(y^*) + \partial h(y^*)$$

KKT系统of（5）