交替方向法的收敛性证明

交替方向法处理的是两个可分离块的凸优化问题

$$\min\{\theta_1(x)+\theta_2(y) \mid Ax+By=b,\ x\in\mathcal{X}, y\in\mathcal{Y}\}. \tag{1}$$

将其拉格朗日函数 $L(x,y,\lambda)=\theta_1(x)+\theta_2(y)-\lambda^T(Ax+By-b)$ 的鞍点归结为等价的变分不等式的解点:

$$w^*\in\Omega,\quad \theta(u)-\theta(u^*)+(w-w^*)^TF(w^*)\ge 0,\quad \forall\, w\in\Omega, \tag{2a}$$

其中 $\Omega=\mathcal{X}\times\mathcal{Y}\times\Re^m$,

$$w=\begin{pmatrix}x\\y\\\lambda\end{pmatrix},\quad u=\begin{pmatrix}x\\y\end{pmatrix},\quad \theta(u)=\theta_1(x)+\theta_2(y),\quad F(w)=\begin{pmatrix}-A^T\lambda\\-B^T\lambda\\Ax+By-b\end{pmatrix}. \tag{2b}$$

ADMM 的 k 步迭代从给定的 $v^k=(y^k,\lambda^k)$ 出发

$$\begin{cases} x^{k+1}\in\arg\min\big\{\theta_1(x)-x^TA^T\lambda^k+\frac{\beta}{2}\|Ax+By^k-b\|^2 \mid x\in\mathcal{X}\big\}, & (3a)\\ y^{k+1}\in\arg\min\big\{\theta_2(y)-y^TB^T\lambda^k+\frac{\beta}{2}\|Ax^{k+1}+By-b\|^2 \mid y\in\mathcal{Y}\big\}, & (3b)\\ \lambda^{k+1}=\ \lambda^k-\beta(Ax^{k+1}+By^{k+1}-b). & (3c)\end{cases}$$

根据最优性引理 1, ADMM k-步迭代满足

$$\begin{cases} x^{k+1}\in\mathcal{X},\ \theta_1(x)-\theta_1(x^{k+1})+(x-x^{k+1})^T\{-A^T\lambda^k+\beta A^T(Ax^{k+1}+By^k-b)\}\ge 0,\ \forall x\in\mathcal{X},\\ y^{k+1}\in\mathcal{Y},\ \theta_2(y)-\theta_2(y^{k+1})+(y-y^{k+1})^T\{-B^T\lambda^k+\beta B^T(Ax^{k+1}+By^{k+1}-b)\}\ge 0,\ \forall y\in\mathcal{Y},\\ \lambda^{k+1}\in\Re^m,\qquad (\lambda-\lambda^{k+1})^T\{(Ax^{k+1}+By^{k+1}-b)+\frac{1}{\beta}(\lambda^{k+1}-\lambda^k)\}\ge 0,\ \forall\lambda\in\Re^m.\end{cases}$$

利用 $\lambda^{k+1}=\ \lambda^k-\beta(Ax^{k+1}+By^{k+1}-b)$ 上面的式子可以整理改写成

$$\begin{cases} x^{k+1}\in\mathcal{X},\quad \theta_1(x)-\theta_1(x^{k+1})+(x-x^{k+1})^T\{-A^T\lambda^{k+1}+\beta A^TB(y^k-y^{k+1})\}\ge 0,\ \forall x\in\mathcal{X}, & (4a)\\ y^{k+1}\in\mathcal{Y},\quad \theta_2(y)-\theta_2(y^{k+1})+(y-y^{k+1})^T\{-B^T\lambda^{k+1}\}\ge 0,\ \forall y\in\mathcal{Y}, & (4b)\\ \lambda^{k+1}\in\Re^m,\quad (\lambda-\lambda^{k+1})^T\{(Ax^{k+1}+By^{k+1}-b)+\frac{1}{\beta}(\lambda^{k+1}-\lambda^k)\}\ge 0,\ \forall\lambda\in\Re^m. & (4c)\end{cases}$$

在 (4b) 的后半部加上和为零的两项, 得到

$$\begin{cases} \underline{\theta_1(x)-\theta_1(x^{k+1})}+(x-x^{k+1})^T\{-A^T\lambda^{k+1}+\beta A^TB(y^k-y^{k+1})\}\ge 0,\\ \underline{\theta_2(y)-\theta_2(y^{k+1})}+(y-y^{k+1})^T\{-B^T\lambda^{k+1}+\underbrace{\beta B^TB(y^k-y^{k+1})+\beta B^TB(y^{k+1}-y^k)}\}\ge 0,\\ (\lambda-\lambda^{k+1})^T\{(Ax^{k+1}+By^{k+1}-b)+\frac{1}{\beta}(\lambda^{k+1}-\lambda^k)\}\ge 0.\end{cases}$$

利用变分不等式 (2), 进行合理整合, 得到

$$\underline{\theta(u)-\theta(u^{k+1})}+(w-w^{k+1})^TF(w^{k+1})$$
$$+\begin{pmatrix}x-x^{k+1}\\y-y^{k+1}\end{pmatrix}^T\beta\begin{pmatrix}A^T\\B^T\end{pmatrix}B(y^k-y^{k+1})+\begin{pmatrix}y-y^{k+1}\\\lambda-\lambda^{k+1}\end{pmatrix}^T\begin{pmatrix}\beta B^TB & 0\\0 & \frac{1}{\beta}I_m\end{pmatrix}\begin{pmatrix}y^{k+1}-y^k\\\lambda^{k+1}-\lambda^k\end{pmatrix}\ge 0.$$

将上式中那个任意的 w, 设成解点 w^* 便有

$$\theta(u^*)-\theta(u^{k+1})+(w^*-w^{k+1})^TF(w^{k+1})$$
$$+\begin{pmatrix} x^*-x^{k+1}\\ y^*-y^{k+1}\end{pmatrix}^T\beta\begin{pmatrix}A^T\\ B^T\end{pmatrix}B(y^k-y^{k+1})+\begin{pmatrix} y^*-y^{k+1}\\ \lambda^*-\lambda^{k+1}\end{pmatrix}^T\begin{pmatrix}\beta B^TB & 0\\ 0 & \frac{1}{\beta}I_m\end{pmatrix}\begin{pmatrix} y^{k+1}-y^k\\ \lambda^{k+1}-\lambda^k\end{pmatrix}\ge 0.$$

经转换, 得到

$$\begin{pmatrix} y^{k+1}-y^*\\ \lambda^{k+1}-\lambda^*\end{pmatrix}^T\begin{pmatrix}\beta B^TB & 0\\ 0 & \frac{1}{\beta}I_m\end{pmatrix}\begin{pmatrix} y^k-y^{k+1}\\ \lambda^k-\lambda^{k+1}\end{pmatrix}\qquad \text{后面记}\ v=\begin{pmatrix}y\\ \lambda\end{pmatrix},\ H=\begin{pmatrix}\beta B^TB & 0\\ 0 & \frac{1}{\beta}I_m\end{pmatrix}$$
$$\ge\begin{pmatrix} x^{k+1}-x^*\\ y^{k+1}-y^*\end{pmatrix}^T\beta\begin{pmatrix}A^T\\ B^T\end{pmatrix}B(y^k-y^{k+1})+\underbrace{[\theta(u^{k+1})-\theta(u^*)+(w^{k+1}-w^*)^TF(w^{k+1})]}. \tag{5}$$

假如 (5) 式右端非负, 证明就基本上完成了. 由于

$$\theta(u^{k+1})-\theta(u^*)+(w^{k+1}-w^*)^TF(w^{k+1})=\theta(u^{k+1})-\theta(u^*)+(w^{k+1}-w^*)^TF(w^*)\ge 0.$$

(5) 式右端下划线部分非负. 因此从 (5) 式得到

$$(v^{k+1}-v^*)^TH(v^k-v^{k+1})\ \ge\ \begin{pmatrix} x^{k+1}-x^*\\ y^{k+1}-y^*\end{pmatrix}^T\beta\begin{pmatrix}A^T\\ B^T\end{pmatrix}B(y^k-y^{k+1}). \tag{6}$$

对 (6) 式的右端进行处理, 有

$$\begin{pmatrix} x^{k+1}-x^*\\ y^{k+1}-y^*\end{pmatrix}^T\beta\begin{pmatrix}A^T\\ B^T\end{pmatrix}B(y^k-y^{k+1})=(y^k-y^{k+1})^TB^T\beta(A,B)\begin{pmatrix} x^{k+1}-x^*\\ y^{k+1}-y^*\end{pmatrix}$$
$$=\ (y^k-y^{k+1})^TB^T\beta\big(Ax^{k+1}+By^{k+1}-(Ax^*+By^*)\big)\qquad \text{利用}\ (Ax^*+By^*=b)$$
$$=\ (y^k-y^{k+1})B^T\beta\big(Ax^{k+1}+By^{k+1}-b\big)=(y^k-y^{k+1})B^T(\lambda^k-\lambda^{k+1}). \tag{7}$$

利用 (4b) 有
$$\theta_2(y)-\theta_2(y^{k+1})\ +\ (y-y^{k+1})^T\{-B^T\lambda^{k+1}\}\ \ge\ 0,\ \forall\, y\in\mathcal{Y},$$
$$\text{和}\ \ \theta_2(y)-\theta_2(y^k)\ +\ (y-y^k)^T\{-B^T\lambda^k\}\ \ge\ 0,\ \forall\, y\in\mathcal{Y}.$$

$$\begin{pmatrix}\text{将任意的}\ y\ \text{分别}\\ \text{设成}\ y^k\ \text{和}\ y^{k+1}\end{pmatrix}\qquad \begin{matrix}\theta_2(y^k)-\theta_2(y^{k+1})\ +\ (y^k-y^{k+1})^T\{-B^T\lambda^{k+1}\}\ \ge 0.\\ \theta_2(y^{k+1})-\theta_2(y^k)\ +\ (y^{k+1}-y^k)^T\{-B^T\lambda^k\}\ \ge 0.\end{matrix}$$

(将上面两式相加, 就有) $\quad (y^k-y^{k+1})B^T(\lambda^k-\lambda^{k+1})\ge 0. \quad$ ((7) 式右端非负)

证明了 (7) 式右端非负, 进而得到 (6) 式右端非负. 所以

$$(v^{k+1}-v^*)^TH(v^k-v^{k+1})\ge 0. \tag{8}$$

Lemma 2 告诉我们:

$$b^TH(a-b)\ge 0\qquad\Rightarrow\qquad \|b\|_H^2\le\|a\|_H^2-\|a-b\|_H^2. \tag{9}$$

在 (9) 中置 $a=(v^k-v^*)$ 和 $b=(v^{k+1}-v^*)$, 根据 (8) 就得到收敛的关键不等式

$$\|v^{k+1}-v^*\|_H^2\le\|v^k-v^*\|_H^2-\|v^k-v^{k+1}\|_H^2.$$