

Properties of saddle points and the partial primal-dual gap

We know that a saddle point (\tilde{u}, \tilde{p}) meets

$$PD(\tilde{u}, p) \leq PD(\tilde{u}, \tilde{p}) \leq PD(u, \tilde{p}) \quad \forall (u, p).$$

Therefore,

$$\begin{aligned} \mathcal{G}_{B_1, B_2}(u, p) &= \max_{p' \in B_2} PD(u, p') - \min_{u' \in B_1} PD(u', p), \\ &\geq PD(u, \tilde{p}) - PD(\tilde{u}, p), \\ &\geq 0, \end{aligned}$$

holds for all (u, p) .

To get some more understanding of saddle-points, note that \tilde{u} minimizes $PD(u, \tilde{p})$ over all u such that

$$-K^* \tilde{p} \in \partial G(\tilde{u}).$$

Similarly \tilde{p} maximizes $PD(\tilde{u}, p)$ over all p such that

$$K \tilde{u} \in \partial F^*(\tilde{p}).$$

Since the latter condition is equivalent to $\tilde{p} \in \partial F(K \tilde{u})$, we find

$$0 = -K^* \tilde{p} + K^* \tilde{p} \in \partial G(\tilde{u}) + K^* \partial F(K \tilde{u}),$$

which means \tilde{u} minimizes $G(u) + F(Ku)$. Similarly, one can show that \tilde{p} maximizes the dual problem $-F^*(p) - G^*(-K^*p)$. Furthermore, any pair (u, p) that has this property is a saddle point. Also note that

$$PD(\tilde{u}, \tilde{p}) = G(\tilde{u}) + \langle K \tilde{u}, \tilde{p} \rangle - F^*(\tilde{p}) = G(\tilde{u}) + F(K \tilde{u}) = -G^*(-K^* \tilde{p}) - F^*(\tilde{p}) = \min_u \max_p PD(u, p).$$

Now let $\mathcal{G}_{B_1, B_2}(v, q) = 0$. In this case

$$\begin{aligned} 0 &= \max_{p' \in B_2} PD(v, p') - \min_{u' \in B_1} PD(u', q), \\ &\geq PD(v, p') - PD(u', q) \quad \forall (u', p') \in (B_1 \times B_2), \end{aligned}$$

which means $PD(v, p') \leq PD(u', q)$ for all $(u', p') \in (B_1 \times B_2)$, or

$$PD(v, p') \leq PD(v, q) \leq PD(u', q), \quad \forall (u', p') \in (B_1 \times B_2).$$

Now since $(\tilde{u}, \tilde{p}) \in (B_1 \times B_2)$, we find

$$PD(\tilde{u}, \tilde{p}) \leq PD(v, \tilde{p}) \leq PD(v, q) \leq PD(\tilde{u}, q) \leq PD(\tilde{u}, \tilde{p})$$

which means that equality holds everywhere, i.e. $PD(v, q) = PD(\tilde{u}, \tilde{p}) = \min_u \max_p PD(u, p)$, which means that (v, q) is a saddle point.

Convergence proof of (PDHG)

We have the update equations

$$\begin{aligned} p^{k+1} &= \arg \min_p \frac{1}{2} \|p - (p^k + \sigma K \bar{u}^k)\|^2 + \sigma F^*(p) \\ u^{k+1} &= \arg \min_u \frac{1}{2} \|u - (u^k - \tau K^* p^{k+1})\|^2 + \tau G(u) \end{aligned}$$

which yield the optimality conditions

$$\begin{aligned} 0 &\in p^{k+1} - p^k - \sigma K \bar{u}^k + \sigma \partial F^*(p^{k+1}), \\ \Rightarrow \frac{p^k - p^{k+1}}{\sigma} + K \bar{u}^k &\in \partial F^*(p^{k+1}), \\ 0 &\in u^{k+1} - u^k + \tau K^* p^{k+1} + \tau \partial G(u^{k+1}), \\ \Rightarrow \frac{u^k - u^{k+1}}{\tau} - \tau K^* p^{k+1} &\in \partial G(u^{k+1}). \end{aligned}$$

By the definition of the subdifferential this yields

$$\begin{aligned} F^*(p) - F^*(p^{k+1}) - \left\langle \frac{p^k - p^{k+1}}{\sigma} + K \bar{u}^k, p - p^{k+1} \right\rangle &\geq 0 \\ G(u) - G(u^{k+1}) - \left\langle \frac{u^k - u^{k+1}}{\tau} - \tau K^* p^{k+1}, u - u^{k+1} \right\rangle &\geq 0 \end{aligned}$$

We add both inequalities to find

$$\begin{aligned} 0 &\geq -F^*(p) + F^*(p^{k+1}) + \left\langle \frac{p^k - p^{k+1}}{\sigma} + K \bar{u}^k, p - p^{k+1} \right\rangle \\ &\quad - G(u) + G(u^{k+1}) + \left\langle \frac{u^k - u^{k+1}}{\tau} - K^* p^{k+1}, u - u^{k+1} \right\rangle \\ &= [G(u^{k+1}) + \langle K u^{k+1}, p \rangle - F^*(p)] - [G(u) + \langle K u, p^{k+1} \rangle - F^*(p^{k+1})] \\ &\quad + \left\langle \frac{p^k - p^{k+1}}{\sigma} + K \bar{u}^k, p - p^{k+1} \right\rangle \\ &\quad + \left\langle \frac{u^k - u^{k+1}}{\tau} - K^* p^{k+1}, u - u^{k+1} \right\rangle \\ &\quad - \langle K u^{k+1}, p \rangle + \langle K u, p^{k+1} \rangle. \end{aligned} \tag{1}$$

We rewrite the remaining inner product terms as

$$\begin{aligned} &\frac{1}{\sigma} \langle p^k - p^{k+1}, p - p^{k+1} \rangle + \langle K \bar{u}^k, p - p^{k+1} \rangle \\ &\quad + \frac{1}{\tau} \langle u^k - u^{k+1}, u - u^{k+1} \rangle - \langle p^{k+1}, K u - K u^{k+1} \rangle \\ &\quad - \langle K u^{k+1}, p \rangle + \langle K u, p^{k+1} \rangle. \\ &= \frac{1}{\sigma} \langle p^k - p^{k+1}, p - p^{k+1} \rangle + \frac{1}{\tau} \langle u^k - u^{k+1}, u - u^{k+1} \rangle \\ &\quad + \langle p - p^{k+1}, K \bar{u}^k - K u^{k+1} \rangle \end{aligned} \tag{2}$$

Now we will use the important identity $2\langle a, b \rangle = \|a\|^2 + \|b\|^2 - \|a - b\|^2$ to see that the above term is

$$\begin{aligned} \dots &= \frac{1}{2\sigma} (\|p^k - p^{k+1}\|^2 + \|p - p^{k+1}\|^2 - \|p^k - p\|) \\ &\quad + \frac{1}{2\tau} (\|u^k - u^{k+1}\|^2 + \|u - u^{k+1}\|^2 - \|u^k - u\|^2) \\ &\quad + \langle p - p^{k+1}, K\bar{u}^k - Ku^{k+1} \rangle \end{aligned} \quad (3)$$

If we think about summing the whole inequality from $k = 1$ to $k = N$, most terms look good already. The only remaining thing to take care of is the last term. At this point we can already see that the choice $\bar{u}^k = u^{k+1}$ would be optimal but not computable and the naive choice $\bar{u}^k = u^k$ would leave us with the problem of how to control this term. For the choice $\bar{u}^k = 2u^k - u^{k-1}$ we can estimate:

$$\begin{aligned} \langle p - p^{k+1}, K\bar{u}^k - Ku^{k+1} \rangle &= \langle p - p^{k+1}, Ku^k - Ku^{k+1} \rangle - \langle p - p^{k+1}, Ku^{k-1} - Ku^k \rangle \\ &= \langle p - p^{k+1}, Ku^k - Ku^{k+1} \rangle - \langle p - p^k, Ku^{k-1} - Ku^k \rangle + \langle p^{k+1} - p^k, Ku^{k-1} - Ku^k \rangle \end{aligned} \quad (4)$$

The remaining thing we have to take care of is the last term, for which we use the Cauchy-Schwartz inequality

$$\begin{aligned} \langle p^{k+1} - p^k, Ku^{k-1} - Ku^k \rangle &\leq \|K\| \|p^{k+1} - p^k\| \|u^{k-1} - u^k\|, \\ &\leq \frac{\|K\|}{2\alpha} \|p^{k+1} - p^k\|^2 + \frac{\alpha\|K\|}{2} \|u^{k-1} - u^k\|^2. \end{aligned} \quad (5)$$

Putting (1), (3), and (5) together yields

$$\begin{aligned} 0 &\geq [G(u^{k+1}) + \langle Ku^{k+1}, p \rangle - F^*(p)] - [G(u) + \langle Ku, p^{k+1} \rangle - F^*(p^{k+1})] \\ &\quad + \frac{1}{2\sigma} (\|p^k - p^{k+1}\|^2 + \|p - p^{k+1}\|^2 - \|p^k - p\|) \\ &\quad + \frac{1}{2\tau} (\|u^k - u^{k+1}\|^2 + \|u - u^{k+1}\|^2 - \|u^k - u\|^2) \\ &\quad + \langle p - p^{k+1}, Ku^k - Ku^{k+1} \rangle - \langle p - p^k, Ku^{k-1} - Ku^k \rangle \\ &\quad - \frac{\|K\|}{2\alpha} \|p^{k+1} - p^k\|^2 - \frac{\alpha\|K\|}{2} \|u^{k-1} - u^k\|^2 \end{aligned} \quad (6)$$

$$\begin{aligned} \Rightarrow 0 &\geq [G(u^{k+1}) + \langle Ku^{k+1}, p \rangle - F^*(p)] - [G(u) + \langle Ku, p^{k+1} \rangle - F^*(p^{k+1})] \\ &\quad + \frac{1}{2\sigma} (\|p - p^{k+1}\|^2 - \|p^k - p\|) + \frac{1}{2\tau} (\|u - u^{k+1}\|^2 - \|u^k - u\|^2) \\ &\quad + \langle p - p^{k+1}, Ku^k - Ku^{k+1} \rangle - \langle p - p^k, Ku^{k-1} - Ku^k \rangle \\ &\quad + \underbrace{\left(\frac{1}{2\sigma} - \frac{\|K\|}{2\alpha} \right)}_{= \frac{1 - \sqrt{\tau\sigma}\|K\|}{2\sigma} > 0} \|p^{k+1} - p^k\|^2 + \frac{1}{2\tau} \|u^{k+1} - u^k\|^2 - \underbrace{\frac{\alpha\|K\|}{2}}_{= \frac{\sqrt{\sigma}\|K\|}{2\sqrt{\tau}}} \|u^{k-1} - u^k\|^2, \end{aligned} \quad (7)$$

for the choice $\alpha = \sqrt{\sigma/\tau}$. Summing over $k = 1$ to $k = N$ yields

$$\begin{aligned} &\sum_{k=1}^N [G(u^{k+1}) + \langle Ku^{k+1}, p \rangle - F^*(p)] - [G(u) + \langle Ku, p^{k+1} \rangle - F^*(p^{k+1})] \\ &\quad + \frac{1}{2\sigma} \|p - p^{N+1}\|^2 + \frac{1}{2\tau} \|u - u^{N+1}\|^2 + \langle p - p^{N+1}, Ku^N - Ku^{N+1} \rangle \\ &\quad + \frac{1 - \sqrt{\tau\sigma}\|K\|}{2\sigma} \sum_{k=1}^N \|p^{k+1} - p^k\|^2 + \frac{1 - \sqrt{\tau\sigma}\|K\|}{2\tau} \sum_{k=1}^N \|u^k - u^{k-1}\|^2 + \frac{1}{2\tau} \|u^{N+1} - u^N\|^2 \\ &\leq \frac{1}{2\sigma} \|p - p^0\|^2 + \frac{1}{2\tau} \|u - u^0\|^2. \end{aligned}$$

Repeating the same estimate as before, we can control the remaining inner product term by the $\|u^{N+1} - u^N\|$ and $\|p - p^{N+1}\|$ terms. The final estimate becomes

$$\begin{aligned}
& \sum_{k=1}^N [G(u^{k+1}) + \langle Ku^{k+1}, p \rangle - F^*(p)] - [G(u) + \langle Ku, p^{p+1} \rangle - F^*(p^{k+1})] \\
& \frac{1 - \sqrt{\tau\sigma}\|K\|}{2\sigma} \sum_{k=1}^N \|p^{k+1} - p^k\|^2 + \frac{1 - \sqrt{\tau\sigma}\|K\|}{2\tau} \sum_{k=1}^N \|u^k - u^{k-1}\|^2 \\
& + \frac{1 - \sqrt{\tau\sigma}\|K\|}{2\sigma} \|p - p^{N+1}\|^2 + \frac{1}{2\tau} \|u - u^{N+1}\|^2 + \frac{1 - \sqrt{\tau\sigma}\|K\|}{2\tau} \|u^{N+1} - u^N\|^2 \\
& \leq \frac{1}{2\sigma} \|p - p^0\|^2 + \frac{1}{2\tau} \|u - u^0\|^2.
\end{aligned}$$

We conclude that the iterates remain bounded, and therefore converge along subsequences. Let (u^{n_k}, p^{n_k}) be a subsequence converging to (u^*, p^*) . Since $\|u^n - u^{n-1}\| \rightarrow 0$ and $\|p^n - p^{n-1}\| \rightarrow 0$, (u^{n_k-1}, p^{n_k-1}) converge to (u^*, p^*) , too. Thus, (u^*, p^*) is a fixed point of the (PDHG) iteration from which we conclude that it is a saddle point of the original problem. By summing (7) from n^k to N for $(u, p) = (u^*, p^*)$ one can deduce the overall convergence to a saddle point.

Jensens inequality yields the estimate

$$[G(u_N) + \langle Ku_N, p \rangle - F^*(p)] - [G(u) + \langle Ku, p_N \rangle - F^*(p_N)] \leq \frac{1}{N} \left(\frac{1}{2\sigma} \|p - p^0\|^2 + \frac{1}{2\tau} \|u - u^0\|^2 \right).$$

for all (u, p) from which one readily shows the second statement of the convergence theorem.