A Nonconvex Proximal Splitting Algorithm under Moreau-Yosida Regularization –Supplementory Material–

Emanuel Laude **Tao Wu** **Daniel Cremers**
Department of Informatics, Technical University of Munich, Germany

A Proofs

A.1 Proof of Lemma 1

Proof. (Statements 1 & 2) To show the lower boundedness of $\mathfrak{L}_\rho(u^{t+1}, z^{t+1}, y^{t+1})$ we rewrite

$$\begin{aligned}\mathfrak{L}_\rho(u^{t+1}, z^{t+1}, y^{t+1}) &= f(z^{t+1}) + g(u^{t+1}) \\ &+ \frac{\rho}{2}\|Au^{t+1} - z^{t+1} - \lambda y^{t+1}\|^2 \\ &+ \frac{1}{2\lambda}\|Au^{t+1} - z^{t+1}\|^2 \\ &- \frac{1}{2\lambda}\|Au^{t+1} - z^{t+1} - \lambda y^{t+1}\|^2.\end{aligned}$$

We define the quadratic penalty

$$Q(u, z) = f(z) + g(u) + \frac{1}{2\lambda}\|Au - z\|^2. \tag{1}$$

Since $\rho > \frac{1}{\lambda}$ we can further bound $\mathfrak{L}_\rho(u^{t+1}, z^{t+1}, y^{t+1})$ from below by $Q(u^{t+1}, z^{t+1})$:

$$\mathfrak{L}_\rho(u^{t+1}, z^{t+1}, y^{t+1}) \geq Q(u^{t+1}, z^{t+1}).$$

We further bound $Q(u^{t+1}, z^{t+1})$:

$$Q(u^{t+1}, z^{t+1}) \geq e_\lambda f(Au^{t+1}) + g(u^{t+1}),$$

which is bounded from below.
(Statement 3) We find an estimate for $\mathfrak{L}_\rho(u^{t+1}, z^t, y^t) - \mathfrak{L}_\rho(u^t, z^t, y^t)$. By the definition of u^{t+1} as the global minimum of $\mathfrak{L}_\rho(\cdot, z^t, y^t) + \frac{1}{2}\|(\cdot) - u^t\|_M^2$ and $M := \frac{1}{\sigma}I - \rho A^\top A$ positive definite for $\sigma\rho\|A\|^2 < 1$, we have the estimate

$$\mathfrak{L}_\rho(u^{t+1}, z^t, y^t) + \frac{1}{2}\|u^{t+1} - u^t\|_M^2 \leq \mathfrak{L}_\rho(u^t, z^t, y^t).$$

We bound $\frac{1}{2}\|u^{t+1} - u^t\|_M^2$,

$$\begin{aligned}\|u^{t+1} - u^t\|_M^2 &= \langle u^{t+1} - u^t, M(u^{t+1} - u^t)\rangle \\ &= \frac{1}{\sigma}\|u^{t+1} - u^t\|^2 - \rho\|Au^{t+1} - Au^t\|^2 \\ &\geq \left(\frac{1}{\sigma} - \rho\|A\|^2\right)\|u^{t+1} - u^t\|^2.\end{aligned}$$

This yields the estimate

$$\begin{aligned}&\mathfrak{L}_\rho(u^{t+1}, z^t, y^t) - \mathfrak{L}_\rho(u^t, z^t, y^t) \\ &\leq \left(\frac{\rho\|A\|^2}{2} - \frac{1}{2\sigma}\right)\|u^{t+1} - u^t\|^2,\end{aligned} \tag{2}$$

which leads to a sufficient descent if $\sigma\rho\|A\|^2 < 1$. The optimality for the z-update guarantees

$$\mathfrak{L}_\rho(u^{t+1}, z^{t+1}, y^t) - \mathfrak{L}_\rho(u^{t+1}, z^t, y^t) \leq 0. \tag{3}$$

Finally we bound the term

$$\begin{aligned}\mathfrak{L}_\rho(u^{t+1}, z^{t+1}, y^{t+1}) - \mathfrak{L}_\rho(u^{t+1}, z^{t+1}, y^t) &= -\frac{\lambda}{2}\|y^{t+1}\|^2 \\ &+ \frac{\lambda}{2}\|y^t\|^2 + \langle Au^{t+1} - z^{t+1}, y^{t+1} - y^t\rangle \\ &+ \frac{\rho}{2}\|Au^{t+1} - z^{t+1} - \lambda y^{t+1}\|^2 \\ &- \frac{\rho}{2}\|Au^{t+1} - z^{t+1} - \lambda y^t\|^2.\end{aligned}$$

Since $\frac{1}{\rho}(y^{t+1} - y^t) + \lambda y^{t+1} = Au^{t+1} - z^{t+1}$, we can rewrite

$$\begin{aligned}&-\frac{\lambda}{2}\|y^{t+1}\|^2 + \frac{\lambda}{2}\|y^t\|^2 + \langle Au^{t+1} - z^{t+1}, y^{t+1} - y^t\rangle \\ &= -\frac{\lambda}{2}\|y^{t+1}\|^2 + \frac{\lambda}{2}\|y^t\|^2 + \frac{1}{\rho}\|y^{t+1} - y^t\|^2 + \lambda\|y^{t+1}\|^2 \\ &\quad - \lambda\langle y^{t+1}, y^t\rangle \\ &= \frac{\lambda}{2}\|y^{t+1}\|^2 - \lambda\langle y^{t+1}, y^t\rangle + \frac{\lambda}{2}\|y^t\|^2 + \frac{1}{\rho}\|y^{t+1} - y^t\|^2 \\ &= \left(\frac{1}{\rho} + \frac{\lambda}{2}\right)\|y^{t+1} - y^t\|^2.\end{aligned}$$

We apply the identity $\|a + c\|^2 - \|b + c\|^2 = -\|b - a\|^2 + 2\langle a + c, a - b\rangle$ with $a := -\lambda y^{t+1}$, $b := -\lambda y^t$ and $c := Au^{t+1} - z^{t+1}$ and obtain

$$\begin{aligned}&\frac{\rho}{2}\|Au^{t+1} - z^{t+1} - \lambda y^{t+1}\|^2 - \frac{\rho}{2}\|Au^{t+1} - z^{t+1} - \lambda y^t\|^2 \\ &= -\frac{\rho\lambda^2}{2}\|y^{t+1} - y^t\|^2 \\ &\quad - \lambda\rho\langle Au^{t+1} - z^{t+1} - \lambda y^{t+1}, y^{t+1} - y^t\rangle \\ &= -\frac{\rho\lambda^2 + 2\lambda}{2}\|y^{t+1} - y^t\|^2.\end{aligned}$$

Overall we have:

$$\begin{aligned}&\mathfrak{L}_\rho(u^{t+1}, z^{t+1}, y^{t+1}) - \mathfrak{L}_\rho(u^{t+1}, z^{t+1}, y^t) \\ &\quad= \left(\frac{1}{\rho} - \frac{\rho\lambda^2 + \lambda}{2}\right) \|y^{t+1} - y^t\|^2.\end{aligned} \tag{4}$$

Summing (2)–(4), we obtain the desired result:

$$\begin{aligned}&\mathcal{Q}_\rho(u^{t+1}, z^{t+1}, y^{t+1}) - \mathcal{Q}_\rho(u^t, z^t, y^t) \\ &\quad\leq \left(\frac{\rho\|A\|^2}{2} - \frac{1}{2\sigma}\right) \|u^{t+1} - u^t\|^2 \\ &\qquad+ \left(\frac{1}{\rho} - \frac{\rho\lambda^2 + \lambda}{2}\right) \|y^{t+1} - y^t\|^2.\end{aligned} \tag{5}$$

□

A.2 Proof of Lemma 2

Proof. Since $\{\mathfrak{L}_\rho(u^t, z^t, y^t)\}_{t\in\mathbb{N}}$ monotonically decreases by Lemma 1, it is bounded from above. Since $\{Q(u^t, z^t)\}_{t\in\mathbb{N}}$ is bounded from above by $\{\mathfrak{L}_\rho(u^t, z^t, y^t)\}_{t\in\mathbb{N}}$ and, furthermore, Q is coercive by assumption, we assert that $\{u^t\}_{t\in\mathbb{N}}$, $\{z^t\}_{t\in\mathbb{N}}$ are uniformly bounded.

Now we sum the estimate (5) from $t = 1$ to T and obtain due to the lower boundedness of the iterates $\mathfrak{L}_\rho(u^t, z^t, y^t)$:

$$\begin{aligned}-\infty &< \mathfrak{L}_\rho(u^{T+1}, z^{T+1}, y^{T+1}) - \mathfrak{L}_\rho(u^1, z^1, y^1) \\ &\leq \left(\frac{\rho\|A\|^2}{2} - \frac{1}{2\sigma}\right) \sum_{t=1}^{T} \|u^{t+1} - u^t\|^2 \\ &\quad+ \left(\frac{1}{\rho} - \frac{\rho\lambda^2 + \lambda}{2}\right) \sum_{t=1}^{T} \|y^{t+1} - y^t\|^2.\end{aligned}$$

Passing $T \to \infty$ yields that $\|u^{t+1} - u^t\| \to 0$ and $\|y^{t+1} - y^t\| \to 0$ for $\rho > 1/\lambda$ and $\sigma\rho\|A\|^2 < 1$. From $\frac{1}{\rho}(y^{t+1} - y^t) = Au^{t+1} - z^{t+1} - \lambda y^{t+1}$ we have that,

$$\begin{aligned}0 &\leq \|z^t - z^{t+1}\| \\ &= \|z^t - z^{t+1} + A(u^{t+1} - u^t) - A(u^{t+1} - u^t) \\ &\qquad + \lambda y^{t+1} - \lambda y^t - \lambda y^{t+1} + \lambda y^t\| \\ &\leq \frac{1}{\rho}\|y^{t+1} - y^t\| + \|A\|\|u^{t+1} - u^t\| \\ &\qquad + \lambda\|y^{t+1} - y^t\| \to 0,\end{aligned}$$

and that $\|Au^t - z^t - \lambda y^t\| \to 0$. Since $\{u^t\}_{t\in\mathbb{N}}$, $\{z^t\}_{t\in\mathbb{N}}$ are uniformly bounded, also $\{y^t\}_{t\in\mathbb{N}}$ are uniformly bounded. □