

## Weekly Exercises 9

Room 01.09.014

Friday, 24.6.2016, 09:00-11:00

Submission deadline: Wednesday, 22.6.2016, 14:15, Room: 02.09.023

### Theory: Monotone Operators (12+6 Points)

**Exercise 1** (4 points). Let  $L \in \mathbb{R}^{n \times n}$  and the operator  $T \subseteq \mathbb{R}^n \times \mathbb{R}^n$  be monotone. Show that

$$L^\top \circ T \circ L$$

is monotone.

**Solution.** Let  $u, v \in \mathbb{R}^n$ . Since  $T$  is monotone we have:

$$\begin{aligned} \langle u - v, (L^\top \circ T \circ L)(u) - (L^\top \circ T \circ L)(v) \rangle &= \langle L(u - v), T(Lu) - T(Lv) \rangle \\ &= \langle Lu - Lv, T(Lu) - T(Lv) \rangle \\ &\geq 0 \end{aligned}$$

**Exercise 2** (4 Points). Let  $E : \mathbb{R}^n \rightarrow \mathbb{R}$  be differentiable. As we have seen in the lecture gradient descent can be formulated in the framework of monotone operators as

$$u^{k+1} = (I - \tau \nabla E)u^k$$

1. Suppose  $E$  is  $m$ -strongly convex and  $L$ -smooth. Show that  $I - \tau \nabla E$  is Lipschitz with

$$L_{GM} = \max\{|1 - \tau m|, |1 - \tau L|\}.$$

2. Let  $E$  be convex and  $L$ -smooth. Show that  $I - \nabla E$  is Lipschitz with

$$L_{GM} = \max\{1, |1 - \tau L|\}.$$

Hint: You may assume that  $E$  is 2 times continuously differentiable.

**Solution.** 1. Let  $u \in \mathbb{R}^n$ . We show the Lipschitz continuity of the mapping  $I - \tau \nabla E$  via giving a bound on its Jacobian (wrt. the operator norm):

$$\nabla(I - \tau \nabla E)(u) = I - \tau \nabla^2 E(u),$$

where the 1st  $I$  refers to the identity mapping and the 2nd  $I$  refers to the identity matrix. The operator norm of  $\|I - \tau \nabla^2 E(u)\|$  is given by the eigenvalue of  $I - \tau \nabla^2 E(u)$  with the largest absolute value. Since

$$m \cdot I \preceq \nabla^2 E(u) \preceq L \cdot I,$$

$m, L$  are the smallest and largest eigenvalues of  $\nabla^2 E(u)$ . The smallest and the largest eigenvalues of  $I - \tau \nabla^2 E(u)$  are the given by  $1 - \tau L$  and  $1 - \tau m$ . Therefore

$$\|I - \tau \nabla^2 E(u)\| \leq \max\{|1 - \tau m|, |1 - \tau L|\}.$$

2. Setting  $m := 0$  and using the above result yields the assertion.

**Exercise 3** (6 Points). Consider the primal-dual hybrid gradient method (PDHG) in operator form:

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} \in \underbrace{\begin{pmatrix} \partial G & -K^* \\ K & \partial F^* \end{pmatrix}}_{=:A} \underbrace{\begin{pmatrix} u^{k+1} \\ p^{k+1} \end{pmatrix}}_{=:z^{k+1}} + \underbrace{\begin{pmatrix} \frac{1}{\tau} I & K^* \\ K & \frac{1}{\sigma} I \end{pmatrix}}_{=:M} \underbrace{\begin{pmatrix} u^{k+1} - u^k \\ p^{k+1} - p^k \end{pmatrix}}_{=:z^{k+1} - z^k}$$

- Assume that a saddle point  $z^* := (u^*, p^*)$ , i.e. an element with  $0 \in Az^*$  exists. Show that the PDHG iterates are Fejer-monotone, i.e.

$$\|z^{k+1} - z^*\|_M \leq \|z^k - z^*\|_M,$$

where  $\|v\|_M^2 = \langle v, Mv \rangle$  for a positive definite matrix  $M$ .

*Hint: Write the above inclusion with respect to  $z_e^{k+1} = z^{k+1} - z^*$  and  $z_e^k = z^k - z^*$ , and take the inner product with  $z_e^{k+1}$ .*

- Assume that  $G$  and  $F^*$  are  $\gamma$ -strongly convex, but  $\tau\sigma\|K\|^2 = 1$  such that  $M$  is positive semi-definite only. Show that  $z^k \rightarrow z^*$ , where  $z^*$  is the (unique) saddle point ( $0 \in Az^*$ ).

*Hint:  $(A - \gamma I)$  is still a monotone operator!*

**Solution.** We define  $z_e^k = z^k - z^*$ . Using  $0 \in Az^*$  we find

$$\begin{aligned} 0 &= \langle z_e^{k+1}, Az_e^{k+1} \rangle + \langle z_e^{k+1}, M(z_e^{k+1} - z_e^k) \rangle \\ &= \langle z_e^{k+1}, Az_e^{k+1} \rangle + \|z_e^{k+1}\|_M^2 - \langle z_e^{k+1}, M - z_e^k \rangle \\ &= \langle z_e^{k+1}, Az_e^{k+1} \rangle + \frac{1}{2} \|z_e^{k+1}\|_M^2 + \frac{1}{2} \|z_e^{k+1}\|_M^2 - \langle z_e^{k+1}, Mz_e^k \rangle + \frac{1}{2} \|z_e^k\|_M^2 - \frac{1}{2} \|z_e^k\|_M^2 \\ &= \langle z_e^{k+1}, Az_e^{k+1} \rangle + \frac{1}{2} \|z_e^{k+1}\|_M^2 + \frac{1}{2} \|z_e^{k+1} - z_e^k\|_M^2 - \frac{1}{2} \|z_e^k\|_M^2 \\ &\geq \frac{1}{2} \|z_e^{k+1}\|_M^2 - \frac{1}{2} \|z_e^k\|_M^2, \end{aligned}$$

where we have used that  $A$  is a monotone operator. This proves the Fejer-monotonicity.

Now let  $G$  and  $F^*$  be  $\gamma$ -strongly convex, which implies that  $A - \gamma I$  is still a monotone operator. Repeating the above computation and noting that  $\|\cdot\|_M$  is a semi-norm only, we find

$$\begin{aligned} 0 &= \langle z_e^{k+1}, Az_e^{k+1} \rangle + \frac{1}{2} \|z_e^{k+1}\|_M^2 + \frac{1}{2} \|z_e^{k+1} - z_e^k\|_M^2 - \frac{1}{2} \|z_e^k\|_M^2 \\ &= \langle z_e^{k+1}, (A - \gamma I)z_e^{k+1} \rangle + \gamma \|z_e^{k+1}\|^2 + \frac{1}{2} \|z_e^{k+1}\|_M^2 + \frac{1}{2} \|z_e^{k+1} - z_e^k\|_M^2 - \frac{1}{2} \|z_e^k\|_M^2 \\ &\geq \gamma \|z_e^{k+1}\|^2 + \frac{1}{2} \|z_e^{k+1}\|_M^2 - \frac{1}{2} \|z_e^k\|_M^2. \end{aligned}$$

Summing over the inequality from  $k = 0$  to  $k = N$  we find

$$\frac{1}{2\gamma} \|z_e^0\|_M^2 \geq \sum_{k=0}^N \|z_e^{k+1}\|^2,$$

which implies  $\|z_e^{k+1}\|^2 \rightarrow 0$ , i.e. convergence of the iterates to the saddle point.

**Warning:** Note that the above way of writing is a slight abuse of notation as  $A$  is a set-valued operator. By the above  $Az^{k+1}$  we mean exactly those elements from the corresponding set which allow to represent 0.

**Exercise 4** (4 Points). Let  $G : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\}$  be proper, closed and convex. Let  $K \in \mathbb{R}^{m \times n}$ . Show the following identity.

$$\text{prox}_{\sigma(G^* \circ -K^*)}(v) = v + \sigma K \cdot \arg \min_u G(u) + \frac{\sigma}{2} \left\| Ku + \frac{v}{\sigma} \right\|^2.$$

**Solution.**

$$\begin{aligned} \hat{u} = \operatorname{argmin}_u G(u) + \frac{\sigma}{2} \left\| Ku + \frac{v}{\sigma} \right\|^2 &\iff 0 \in \partial G(\hat{u}) + \sigma K^* \left( K\hat{u} + \frac{v}{\sigma} \right) \\ &\iff -K^*(v + \sigma K\hat{u}) \in \partial G(\hat{u}) \\ &\iff \hat{u} \in \partial G^*(-K^*(v + \sigma K\hat{u})) \\ &\implies v + \sigma K\hat{u} \in v + \sigma K \partial G^*(-K^*(v + \sigma K\hat{u})) \\ &\iff v + \sigma K\hat{u} \in v - \sigma \partial(G^* \circ -K^*)(v + \sigma K\hat{u}) \\ &\iff 0 \in (v + \sigma K\hat{u}) - v + \sigma \partial(G^* \circ -K^*)(v + \sigma K\hat{u}) \\ &\iff 0 \in \frac{1}{\sigma} ((v + \sigma K\hat{u}) - v) + \partial(G^* \circ -K^*)(v + \sigma K\hat{u}) \\ &\iff v + \sigma K\hat{u} = \operatorname{argmin}_u \frac{1}{2\sigma} \|u - v\|_2^2 + (G^* \circ -K^*)(u) \\ &\iff v + \sigma K\hat{u} = \text{prox}_{\sigma(G^* \circ -K^*)}(v) \end{aligned}$$

## Programming: Deblurring revisited (12 Points)

**Exercise 5** (12 Points). Given a blurry and noisy input image  $f$ , reconstruct a sharper image  $u^*$  by solving the following optimization problem

$$u^* = \arg \min_u \frac{1}{2} \|k * u - f\|^2 + \alpha \|Du\|_{2,1}, \quad (1)$$

with ADMM. For details on how to deal with the convolution cf. sheet 6. Since we consider the non-blind setting, i.e. a known blur kernel  $k$ , we obtain a blurry and noisy input image by convolving the image with the kernel  $k$  and adding a small amount of Gaussian noise.