

Dual and primal-dual methods

Ye Shi

ShanghaiTech University

Outline

- Dual proximal gradient method
- Primal-dual proximal gradient method

Dual proximal gradient method

Constrained convex optimization

$$\begin{array}{ll}\text{minimize}_x & f(x) \\ \text{subject to} & \mathbf{A}x + \mathbf{b} \in \mathcal{C}\end{array}$$

where f is convex, and \mathcal{C} is convex set

- projection onto such a feasible set could sometimes be highly nontrivial (even when projection onto \mathcal{C} is easy)

Constrained convex optimization

More generally, consider

$$\text{minimize}_x \quad f(\mathbf{x}) + h(\mathbf{Ax})$$

where f and h are convex

- computing the proximal operator w.r.t. $\tilde{h}(\mathbf{x}) := h(\mathbf{Ax})$ could be difficult (even when prox_h is inexpensive)

A possible route: dual formulation

$$\text{minimize}_x \quad f(x) + h(Ax)$$

\Updownarrow add auxiliary variable z

$$\begin{aligned} \text{minimize}_{x,z} \quad & f(x) + h(z) \\ \text{subject to} \quad & Ax = z \end{aligned}$$

dual formulation:

$$\text{maximize}_\lambda \quad \min_{x,z} \underbrace{f(x) + h(z) + \langle \lambda, Ax - z \rangle}_{=:\mathcal{L}(x,z,\lambda) \text{ (Lagrangian)}}$$

A possible route: dual formulation

$$\text{maximize}_{\lambda} \quad \min_{\mathbf{x}, \mathbf{z}} f(\mathbf{x}) + h(\mathbf{z}) + \langle \lambda, \mathbf{Ax} - \mathbf{z} \rangle$$

\Updownarrow decouple \mathbf{x} and \mathbf{z}

$$\text{maximize}_{\lambda} \quad \min_{\mathbf{x}} \left\{ \langle \mathbf{A}^{\top} \lambda, \mathbf{x} \rangle + f(\mathbf{x}) \right\} + \min_{\mathbf{z}} \left\{ h(\mathbf{z}) - \langle \lambda, \mathbf{z} \rangle \right\}$$

\Updownarrow

$$\text{maximize}_{\lambda} \quad -f^*(-\mathbf{A}^{\top} \lambda) - h^*(\lambda)$$

where f^* (resp. h^*) is the Fenchel conjugate of f (resp. h)

Primal vs. dual problems

$$\text{(primal)} \quad \text{minimize}_{\mathbf{x}} \quad f(\mathbf{x}) + h(\mathbf{A}\mathbf{x})$$

$$\text{(dual)} \quad \text{minimize}_{\boldsymbol{\lambda}} \quad f^*(-\mathbf{A}^\top \boldsymbol{\lambda}) + h^*(\boldsymbol{\lambda})$$

Dual formulation is useful if

- the proximal operator w.r.t. h is cheap (then we can use the Moreau decomposition $\text{prox}_{h^*}(\mathbf{x}) = \mathbf{x} - \text{prox}_h(\mathbf{x})$)
- f^* is smooth (or if f is strongly convex)

Dual proximal gradient methods

Apply proximal gradient methods to the dual problem:

Algorithm 9.1 Dual proximal gradient algorithm

1: **for** $t = 0, 1, \dots$ **do**

2: $\lambda^{t+1} = \text{prox}_{\eta_t h^*} \left(\lambda^t + \eta_t A \nabla f^* (-A^\top \lambda^t) \right)$

- let $Q(\lambda) := -f^*(-A^\top \lambda) - h^*(\lambda)$ and $Q^{\text{opt}} = \max_{\lambda} Q(\lambda)$, then

$$Q^{\text{opt}} - Q(\lambda^t) \lesssim \frac{1}{t} \quad (9.1)$$

Primal representation of dual proximal gradient methods

Algorithm 9.1 admits a more explicit primal representation

Algorithm 9.2 Dual proximal gradient algorithm (primal representation)

```
1: for  $t = 0, 1, \dots$  do  
2:    $\mathbf{x}^t = \arg \min_{\mathbf{x}} \{f(\mathbf{x}) + \langle \mathbf{A}^\top \boldsymbol{\lambda}^t, \mathbf{x} \rangle\}$   
3:    $\boldsymbol{\lambda}^{t+1} = \boldsymbol{\lambda}^t + \eta_t \mathbf{A} \mathbf{x}^t - \eta_t \text{prox}_{\eta_t^{-1}h}(\eta_t^{-1} \boldsymbol{\lambda}^t + \mathbf{A} \mathbf{x}^t)$ 
```

- $\{\mathbf{x}^t\}$ is a primal sequence, which is nonetheless *not always feasible*

Justification of the primal representation

By definition of \mathbf{x}^t ,

$$-\mathbf{A}^\top \boldsymbol{\lambda}^t \in \partial f(\mathbf{x}^t)$$

This together with the conjugate subgradient theorem and the smoothness of f^* yields

$$\mathbf{x}^t = \nabla f^*(-\mathbf{A}^\top \boldsymbol{\lambda}^t)$$

Therefore, the dual proximal gradient update rule can be rewritten as

$$\boldsymbol{\lambda}^{t+1} = \text{prox}_{\eta_t h^*}(\boldsymbol{\lambda}^t + \eta_t \mathbf{A} \mathbf{x}^t) \tag{9.2}$$

Justification of primal representation (cont.)

Moreover, from the extended Moreau decomposition, we know

$$\begin{aligned}\operatorname{prox}_{\eta_t h^*}(\boldsymbol{\lambda}^t + \eta_t \mathbf{A} \mathbf{x}^t) &= \boldsymbol{\lambda}^t + \eta_t \mathbf{A} \mathbf{x}^t - \eta_t \operatorname{prox}_{\eta_t^{-1} h}(\eta_t^{-1} \boldsymbol{\lambda}^t + \mathbf{A} \mathbf{x}^t) \\ \implies \boldsymbol{\lambda}^{t+1} &= \boldsymbol{\lambda}^t + \eta_t \mathbf{A} \mathbf{x}^t - \eta_t \operatorname{prox}_{\eta_t^{-1} h}(\eta_t^{-1} \boldsymbol{\lambda}^t + \mathbf{A} \mathbf{x}^t)\end{aligned}$$

Accuracy of the primal sequence

One can control the primal accuracy via the dual accuracy:

Lemma 9.1

Let $\mathbf{x}_\lambda := \arg \min_{\mathbf{x}} \{f(\mathbf{x}) + \langle \mathbf{A}^\top \boldsymbol{\lambda}, \mathbf{x} \rangle\}$. Suppose f is μ -strongly convex. Then

$$\|\mathbf{x}^* - \mathbf{x}_\lambda\|_2^2 \leq \frac{2(Q^{\text{opt}} - Q(\boldsymbol{\lambda}))}{\mu}$$

- **consequence:** $\|\mathbf{x}^* - \mathbf{x}^t\|_2^2 \lesssim 1/t$ (using (9.1))

Proof of Lemma 9.1

Recall that Lagrangian is given by

$$\mathcal{L}(x, z, \lambda) := \underbrace{f(x) + \langle A^\top \lambda, x \rangle}_{=: \tilde{f}(x, \lambda)} + \underbrace{h(z) - \langle \lambda, z \rangle}_{=: \tilde{h}(z, \lambda)}$$

For any λ , define $x_\lambda := \arg \min_x \tilde{f}(x, \lambda)$ and $z_\lambda := \arg \min_z \tilde{h}(z, \lambda)$ (non-rigorous). Then by strong convexity,

$$\mathcal{L}(x^*, z^*, \lambda) - \mathcal{L}(x_\lambda, z_\lambda, \lambda) \geq \tilde{f}(x^*, \lambda) - \tilde{f}(x_\lambda, \lambda) \geq \frac{1}{2} \mu \|x^* - x_\lambda\|_2^2$$

In addition, since $Ax^* = z^*$, one has

$$\begin{aligned} \mathcal{L}(x^*, z^*, \lambda) &= f(x^*) + h(z^*) + \langle \lambda, Ax^* - z^* \rangle = f(x^*) + h(Ax^*) \\ &= F^{\text{opt}} \stackrel{\text{duality}}{=} Q^{\text{opt}} \end{aligned}$$

This combined with $\mathcal{L}(x_\lambda, z_\lambda, \lambda) = Q(\lambda)$ gives

$$Q^{\text{opt}} - Q(\lambda) \geq \frac{1}{2} \mu \|x^* - x_\lambda\|_2^2$$

as claimed

Accelerated dual proximal gradient methods

One can apply FISTA to dual problem to improve convergence:

Algorithm 9.3 Accelerated dual proximal gradient algorithm

- 1: **for** $t = 0, 1, \dots$ **do**
 - 2: $\lambda^{t+1} = \text{prox}_{\eta_t h^*} \left(w^t + \eta_t \mathbf{A} \nabla f^* (-\mathbf{A}^\top w^t) \right)$
 - 3: $\theta_{t+1} = \frac{1 + \sqrt{1 + 4\theta_t^2}}{2}$
 - 4: $w^{t+1} = \lambda^{t+1} + \frac{\theta_t - 1}{\theta_{t+1}} (\lambda^{t+1} - \lambda^t)$
-

- apply FISTA theory and Lemma 9.1 to get

$$Q^{\text{opt}} - Q(\lambda^t) \lesssim \frac{1}{t^2} \quad \text{and} \quad \|x^* - x^t\|_2^2 \lesssim \frac{1}{t^2}$$

Primal representation of accelerated dual proximal gradient methods

Algorithm 9.3 admits more explicit primal representation

Algorithm 9.4 Accelerated dual proximal gradient algorithm (primal representation)

- 1: **for** $t = 0, 1, \dots$ **do**
 - 2: $\mathbf{x}^t = \arg \min_{\mathbf{x}} f(\mathbf{x}) + \langle \mathbf{A}^\top \mathbf{w}^t, \mathbf{x} \rangle$
 - 3: $\boldsymbol{\lambda}^{t+1} = \mathbf{w}^t + \eta_t \mathbf{A} \mathbf{x}^t - \eta_t \text{prox}_{\eta_t^{-1}h}(\eta_t^{-1} \mathbf{w}^t + \mathbf{A} \mathbf{x}^t)$
 - 4: $\theta_{t+1} = \frac{1 + \sqrt{1 + 4\theta_t^2}}{2}$
 - 5: $\mathbf{w}^{t+1} = \boldsymbol{\lambda}^{t+1} + \frac{\theta_t - 1}{\theta_{t+1}} (\boldsymbol{\lambda}^{t+1} - \boldsymbol{\lambda}^t)$
-

Primal-dual proximal gradient method

Nonsmooth optimization

$$\text{minimize}_{\boldsymbol{x}} \quad f(\boldsymbol{x}) + h(\boldsymbol{A}\boldsymbol{x})$$

where f and h are closed and convex

- both f and h might be non-smooth
- both f and h might have inexpensive proximal operators

Primal-dual approaches?

$$\text{minimize}_{\boldsymbol{x}} \quad f(\boldsymbol{x}) + h(\boldsymbol{A}\boldsymbol{x})$$

So far we have discussed proximal methods (resp. dual proximal methods), which essentially updates only primal (resp. dual) variables

Question: can we update both primal and dual variables simultaneously and take advantage of both prox_f and prox_h ?

A saddle-point formulation

To this end, we first derive a saddle-point formulation that includes both primal and dual variables

$$\text{minimize}_x \quad f(x) + h(Ax)$$

\Updownarrow add an auxiliary variable z

$$\text{minimize}_{x,z} \quad f(x) + h(z) \quad \text{subject to } Ax = z$$

\Updownarrow

$$\text{maximize}_\lambda \min_{x,z} f(x) + h(z) + \langle \lambda, Ax - z \rangle$$

\Updownarrow

$$\text{maximize}_\lambda \min_x f(x) + \langle \lambda, Ax \rangle - h^*(\lambda)$$

\Updownarrow

$$\text{minimize}_x \max_\lambda f(x) + \langle \lambda, Ax \rangle - h^*(\lambda) \quad (\text{saddle-point problem})$$

A saddle-point formulation

$$\text{minimize}_x \max_{\lambda} f(x) + \langle \lambda, Ax \rangle - h^*(\lambda) \quad (9.3)$$

- one can then consider updating the primal variable x and the dual variable λ simultaneously
- we'll first examine the optimality condition for (9.3), which in turn gives ideas about how to jointly update primal and dual variables

Optimality condition

$$\text{minimize}_x \max_{\lambda} f(x) + \langle \lambda, Ax \rangle - h^*(\lambda)$$

optimality condition:

$$\begin{cases} \mathbf{0} \in \partial f(x) + A^\top \lambda \\ \mathbf{0} \in -Ax + \partial h^*(\lambda) \end{cases}$$

$$\iff \mathbf{0} \in \begin{bmatrix} & A^\top \\ -A & \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} + \begin{bmatrix} \partial f(x) \\ \partial h^*(\lambda) \end{bmatrix} =: \mathcal{F}(x, \lambda) \quad (9.4)$$

key idea: iteratively update (x, λ) to reach a point obeying $\mathbf{0} \in \mathcal{F}(x, \lambda)$

How to solve $0 \in \mathcal{F}(x)$ in general?

In general, finding solution to

$$\underbrace{0 \in \mathcal{F}(x)}$$

called “monotone inclusion problem” if \mathcal{F} is maximal monotone

$$\iff x \in (\mathcal{I} + \mathcal{F})(x)$$

is equivalent to finding fixed points of $\underbrace{(\mathcal{I} + \eta\mathcal{F})^{-1}}_{\text{resolvent of } \mathcal{F}}$, i.e. solutions to

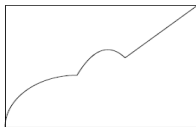
$$x = (\mathcal{I} + \eta\mathcal{F})^{-1}(x)$$

This suggests a natural fixed-point iteration / resolvent iteration:

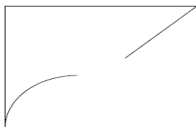
$$x^{t+1} = (\mathcal{I} + \eta\mathcal{F})^{-1}(x^t), \quad t = 0, 1, \dots$$

Aside: monotone operators

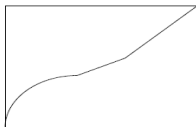
— Ryu, Boyd '16



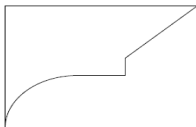
(A) Not monotone.



(B) Monotone but not maximal.



(C) Maximal monotone function.



(D) Maximal monotone but not a function.

- a relation \mathcal{F} is called *monotone* if

$$\langle \mathbf{u} - \mathbf{v}, \mathbf{x} - \mathbf{y} \rangle \geq 0, \quad \forall (\mathbf{x}, \mathbf{u}), (\mathbf{y}, \mathbf{v}) \in \mathcal{F}$$

- relation \mathcal{F} is called *maximal monotone* if there is no monotone operator that contains it

Proximal point method

$$\mathbf{x}^{t+1} = (\mathcal{I} + \eta_t \mathcal{F})^{-1}(\mathbf{x}^t), \quad t = 0, 1, \dots$$

If $\mathcal{F} = \partial f$ for some convex function f , then this **proximal point method** becomes

$$\mathbf{x}^{t+1} = \text{prox}_{\eta_t f}(\mathbf{x}^t), \quad t = 0, 1, \dots$$

- useful when $\text{prox}_{\eta_t f}$ is cheap

Back to primal-dual approaches

Recall that we want to solve

$$\mathbf{0} \in \begin{bmatrix} & \mathbf{A}^\top \\ -\mathbf{A} & \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \boldsymbol{\lambda} \end{bmatrix} + \begin{bmatrix} \partial f(\mathbf{x}) \\ \partial h^*(\boldsymbol{\lambda}) \end{bmatrix} =: \mathcal{F}(\mathbf{x}, \boldsymbol{\lambda})$$

the issue of proximal point methods: computing $(\mathcal{I} + \eta\mathcal{F})^{-1}$ is in general difficult

Back to primal-dual approaches

observation: practically we may often consider splitting \mathcal{F} into two operators

$$\mathbf{0} \in \mathcal{A}(\mathbf{x}, \boldsymbol{\lambda}) + \mathcal{B}(\mathbf{x}, \boldsymbol{\lambda})$$

$$\text{with } \mathcal{A}(\mathbf{x}, \boldsymbol{\lambda}) = \begin{bmatrix} & \mathbf{A} \\ -\mathbf{A}^\top & \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \boldsymbol{\lambda} \end{bmatrix}, \quad \mathcal{B}(\mathbf{x}, \boldsymbol{\lambda}) = \begin{bmatrix} \partial f(\mathbf{x}) \\ \partial h^*(\boldsymbol{\lambda}) \end{bmatrix} \quad (9.5)$$

- $(\mathcal{I} + \eta\mathcal{A})^{-1}$ can be computed by solving linear systems
- $(\mathcal{I} + \eta\mathcal{B})^{-1}$ is easy if prox_f and prox_{h^*} are both inexpensive

solution: design update rules based on $(\mathcal{I} + \eta\mathcal{A})^{-1}$ and $(\mathcal{I} + \eta\mathcal{B})^{-1}$ instead of $(\mathcal{I} + \eta\mathcal{F})^{-1}$

Operator splitting via Cayley operators

We now introduce a principled approach based on operator splitting

$$\text{find } x \quad \text{s.t. } 0 \in \mathcal{F}(x) = \underbrace{\mathcal{A}(x) + \mathcal{B}(x)}_{\text{operator splitting}}$$

let $\mathcal{R}_{\mathcal{A}} := (\mathcal{I} + \eta\mathcal{A})^{-1}$ and $\mathcal{R}_{\mathcal{B}} := (\mathcal{I} + \eta\mathcal{B})^{-1}$ be the **resolvents**, and $\mathcal{C}_{\mathcal{A}} := 2\mathcal{R}_{\mathcal{A}} - \mathcal{I}$ and $\mathcal{C}_{\mathcal{B}} := 2\mathcal{R}_{\mathcal{B}} - \mathcal{I}$ be the **Cayley operators**

Lemma 9.2

$$\underbrace{0 \in \mathcal{A}(x) + \mathcal{B}(x)}_{x \in \mathcal{R}_{\mathcal{A}+\mathcal{B}}(x)} \iff \underbrace{\mathcal{C}_{\mathcal{A}}\mathcal{C}_{\mathcal{B}}(z) = z \text{ with } x = \mathcal{R}_{\mathcal{B}}(z)}_{\text{it comes down to finding fixed points of } \mathcal{C}_{\mathcal{A}}\mathcal{C}_{\mathcal{B}}} \quad (9.6)$$

Operator splitting via Cayley operators

$$x \in \mathcal{R}_{\mathcal{A}+\mathcal{B}}(x) \iff \mathcal{C}_{\mathcal{A}}\mathcal{C}_{\mathcal{B}}(z) = z$$

- **advantage:** allows us to apply $\mathcal{C}_{\mathcal{A}}$ (resp. $\mathcal{R}_{\mathcal{A}}$) and $\mathcal{C}_{\mathcal{B}}$ (resp. $\mathcal{R}_{\mathcal{B}}$) sequentially (instead of computing $\mathcal{R}_{\mathcal{A}+\mathcal{B}}$ directly)

Proof of Lemma 9.2

$$\mathcal{C}_A \mathcal{C}_B(z) = z$$

$$x = \mathcal{R}_B(z) \tag{9.7a}$$

$$\iff \tilde{z} = 2x - z \tag{9.7b}$$

$$\tilde{x} = \mathcal{R}_A(\tilde{z}) \tag{9.7c}$$

$$z = 2\tilde{x} - \tilde{z} \tag{9.7d}$$

From (9.7b) and (9.7d), we see that

$$\tilde{x} = x$$

which together with (9.7d) gives

$$2x = z + \tilde{z} \tag{9.8}$$

Proof of Lemma 9.2 (cont.)

Recall that

$$z \in x + \eta\mathcal{B}(x) \quad \text{and} \quad \tilde{z} \in x + \eta\mathcal{A}(x)$$

Adding these two facts and using (9.8), we get

$$\begin{aligned} 2x = z + \tilde{z} &\in 2x + \eta\mathcal{B}(x) + \eta\mathcal{A}(x) \\ \iff 0 &\in \mathcal{A}(x) + \mathcal{B}(x) \end{aligned}$$

Douglas-Rachford splitting

How to find points obeying $x = \mathcal{C}_A \mathcal{C}_B(x)$?

- First attempt: fixed-point iteration

$$z^{t+1} = \mathcal{C}_A \mathcal{C}_B(z^t)$$

unfortunately, it may not converge in general

- **Douglas-Rachford splitting:** damped fixed-point iteration

$$z^{t+1} = \frac{1}{2}(\mathcal{I} + \mathcal{C}_A \mathcal{C}_B)(z^t)$$

converges when a solution to $0 \in \mathcal{A}(x) + \mathcal{B}(x)$ exists!

More explicit expression for D-R splitting

Douglas-Rachford splitting update rule $\mathbf{z}^{t+1} = \frac{1}{2}(\mathcal{I} + \mathcal{C}_\mathcal{A}\mathcal{C}_\mathcal{B})(\mathbf{z}^t)$ is essentially:

$$\mathbf{x}^{t+\frac{1}{2}} = \mathcal{R}_\mathcal{B}(\mathbf{z}^t)$$

$$\mathbf{z}^{t+\frac{1}{2}} = 2\mathbf{x}^{t+\frac{1}{2}} - \mathbf{z}^t$$

$$\mathbf{x}^{t+1} = \mathcal{R}_\mathcal{A}(\mathbf{z}^{t+\frac{1}{2}})$$

$$\begin{aligned}\mathbf{z}^{t+1} &= \frac{1}{2}(\mathbf{z}^t + 2\mathbf{x}^{t+1} - \mathbf{z}^{t+\frac{1}{2}}) \\ &= \mathbf{z}^t + \mathbf{x}^{t+1} - \mathbf{x}^{t+\frac{1}{2}}\end{aligned}$$

where $\mathbf{x}^{t+\frac{1}{2}}$ and $\mathbf{z}^{t+\frac{1}{2}}$ are auxiliary variables

More explicit expression for D-R splitting

or equivalently,

$$\mathbf{x}^{t+\frac{1}{2}} = \mathcal{R}_{\mathcal{B}}(\mathbf{z}^t)$$

$$\mathbf{x}^{t+1} = \mathcal{R}_{\mathcal{A}}(2\mathbf{x}^{t+\frac{1}{2}} - \mathbf{z}^t)$$

$$\mathbf{z}^{t+1} = \mathbf{z}^t + \mathbf{x}^{t+1} - \mathbf{x}^{t+\frac{1}{2}}$$

Douglas-Rachford primal-dual splitting

$$\text{minimize}_x \max_{\lambda} f(x) + \langle \lambda, Ax \rangle - h^*(\lambda)$$

Applying Douglas-Rachford splitting to (9.5) yields

$$x^{t+\frac{1}{2}} = \text{prox}_{\eta f}(p^t)$$

$$\lambda^{t+\frac{1}{2}} = \text{prox}_{\eta h^*}(q^t)$$

$$\begin{bmatrix} x^{t+1} \\ \lambda^{t+1} \end{bmatrix} = \begin{bmatrix} I & \eta A^\top \\ -\eta A & I \end{bmatrix}^{-1} \begin{bmatrix} 2x^{t+\frac{1}{2}} - p^t \\ 2\lambda^{t+\frac{1}{2}} - q^t \end{bmatrix}$$

$$p^{t+1} = p^t + x^{t+1} - x^{t+\frac{1}{2}}$$

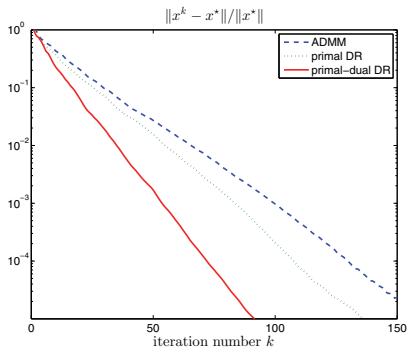
$$q^{t+1} = q^t + \lambda^{t+1} - \lambda^{t+\frac{1}{2}}$$

Example

$$\text{minimize}_x \quad \|x\|_2 + \gamma \|Ax - b\|_1$$

$$\iff \text{minimize}_x \quad f(x) + g(Ax)$$

with $f(x) := \|x\|_2$ and $g(y) := \gamma \|y - b\|_1$



— Connor, Vandenberghe '14

Reference

- [1] "*A First Course in Convex Optimization Theory*," E. Ryu, W. Yin.
- [2] "*Optimization methods for large-scale systems, EE236C lecture notes*," L. Vandenberghe, UCLA.
- [3] "*Convex optimization, EE364B lecture notes*," S. Boyd, Stanford.
- [4] "*Mathematical optimization, MATH301 lecture notes*," E. Candes, Stanford.
- [5] "*First-order methods in optimization*," A. Beck, Vol. 25, SIAM, 2017.
- [6] "*Primal-dual decomposition by operator splitting and applications to image deblurring*," D. O' Connor, L. Vandenberghe, *SIAM Journal on Imaging Sciences*, 2014.
- [7] "*On the numerical solution of heat conduction problems in two and three space variables*," J. Douglas, H. Rachford, *Transactions of the American mathematical Society*, 1956.

Reference

- [8] "*A primer on monotone operator methods*," E. Ryu, S. Boyd, *Appl. Comput. Math.*, 2016.