Topics in Mathematical Imaging
Lecture 3

Carola-Bibiane Schönlieb

Department for Applied Mathematics and Theoretical Physics
Cantab Capital Institute for the Mathematics of Information
EPSRC Centre for Mathematical Imaging in Healthcare
Alan Turing Institute
University of Cambridge, UK

Spring School SFB 1294, March 2018

Lecture plan

- Lecture 1: Variational models & PDEs for imaging by examples
- Lecture 2: Derivation of these models & analysis
- Lecture 3: Numerical solution
- Lecture 4: Some machine learning connections

Setting – discrete

Consider for $u \in \mathbb{R}^n$

$$\min \mathcal{J}(u) + \mathcal{H}(u),$$

where $\mathcal{J}$ and $\mathcal{H}$ are proper and convex, and (possibly) $\mathcal{J}$ and / or $\mathcal{H}$ is Lipschitz differentiable.

Example: ROF problem for $g \in \mathbb{R}^n$ solve

$$\min_u \alpha \|Du\|_{2,1} + \frac{1}{2}\|u - g\|_2^2$$

A quick overview of main approaches for minimising such functionals …

Reference for this part: Chambolle, Pock, Acta Numerica 2016.

Classical methods for smooth problems

…algorithms which attempt to compute minimisers of the regularised ROF problem

$$\min_u \left\{ \alpha \sum \sqrt{u_x^2 + u_y^2 + \epsilon} + \frac{1}{2} \|u - g\|_2^2 \right\}$$

for a small $0 < \epsilon \ll 1$.

Since in this case the regularised TV is differentiable in the classical sense we can apply classical numerical algorithms to compute a minimiser, e.g. gradient descent, conjugate gradient etc.

In what follows: convex algorithms which look at the non-regularised problem.

Preliminary concepts . . .

Subdifferential

UNIVERSITY OF CAMBRIDGE

Definition

For a locally convex space V and for a convex function $F : V \to \mathbb{R} \cup \{-\infty, +\infty\}$, we define the *subdifferential* of F at $x \in V$, as $\partial F(x) = \emptyset$ if $F(x) = \infty$, otherwise

$$\partial F(x) := \partial F_V(x) := \{x^* \in V' : \langle x^*, y - x \rangle + F(x) \leq F(y) \quad \forall y \in V\},$$

where V' denotes the dual space of V. It is obvious from this definition that $0 \in \partial F(x)$ if and only if x is a minimizer of F. We write $\partial_V F$ for the subdifferential considered on the space V.

Example: Let $V = \ell_1(\Lambda)$ and $F(x) := \|x\|_1$ is the ℓ_1-norm. We have

$$\partial \| \cdot \|_1(x) \; = \; \{\xi \in \ell_\infty(\Lambda) : \; \xi_\lambda \in \partial | \cdot |(x_\lambda), \lambda \in \Lambda\} \tag{1}$$

where $\partial | \cdot |(z) = \{ \operatorname{sign}(z)\}$ if $z \neq 0$ and $\partial | \cdot |(0) = [-1, 1]$.

The Legendre-Fenchel transform

For J being one-homogeneous

$$\text{that is, } J(\lambda u) = \lambda J(u) \text{ for every } u \text{ and } \lambda > 0,$$

it is a standard fact in convex analysis that the Legendre-Fenchel transform

$$\text{that is } J^*(v) = \sup_u \langle u, v \rangle_X - J(u) \text{ (with } \langle u, v \rangle_X = \sum_{i,j} u_{i,j} v_{i,j})$$

is the characteristic function of a closed convex set K:

$$J^*(v) = \chi_K(v) = \begin{cases} 0 & \text{if } v \in K \\ +\infty & \text{otherwise.} \end{cases}$$

Since $J^{**} = J$, we recover

$$J(u) = \sup_{v \in K} \langle u, v \rangle_X \,.$$

Proximal map

Let $\mathcal{J}$ convex, proper and l.s.c., then for any f there is a unique minimiser

$$u^* = \text{argmin}_u \mathcal{J}(u) + \frac{1}{2\tau}\|u - f\|_2^2$$

We call $u^* = \text{prox}_{\tau\mathcal{J}}(f)$ the proximal map of $\mathcal{J}$ at f. With optimality condition

$$0 \in \partial\mathcal{J}(u^*) + \frac{u^* - f}{\tau}$$

this reads

$$u^* = (I + \tau\partial J)^{-1} f.$$

Rockafellar 1997

Moreau's identity

UNIVERSITY OF CAMBRIDGE

One can show

$$f = \operatorname{prox}_{\tau\mathcal{J}}(f) + \tau \operatorname{prox}_{\frac{1}{\tau}\mathcal{J}^*}(\frac{f}{\tau}),$$

which shows:

If we know how to compute $\operatorname{prox}_{\mathcal{J}}$ we also know how to compute $\operatorname{prox}_{\mathcal{J}^*}$.

Convex duality

Consider

$$\min_{u \in X} \mathcal{J}(Ku) + \mathcal{H}(u),$$

where $\mathcal{J} : Y \to (-\infty, +\infty]$, $\mathcal{H} : X \to (-\infty, +\infty]$ convex, l.s.c., $K : X \to Y$ linear and bounded. Then (under mild appropriate assumptions on $\mathcal{J}, \mathcal{H}$)

$$\begin{aligned}
&\min_{u \in X} \mathcal{J}(Ku) + \mathcal{H}(u) \\
&\underbrace{=}_{\mathcal{J}^{**}=\mathcal{J}} \min_{u \in X} \sup_{p \in Y} \langle p, Ku \rangle - \mathcal{J}^*(p) + \mathcal{H}(u) \\
&= \max_{p} \inf_{u} \langle p, Ku \rangle - \mathcal{J}^*(p) + \mathcal{H}(u) \\
&= \max_{p} -\mathcal{J}^*(p) - \mathcal{H}^*(-K^* p).
\end{aligned}$$

The latter is the dual problem. Under above assumptions there exists at least one solution p^*. Book, Ekeland, Temam 1999; Survey article by Borwein, Luke 2015

Saddle-point problem

If u^* solves primal problem and p^* dual problem, then (u^*, p^*) is a saddle-point of primal-dual problem

$$\forall (u,p) \in X \times Y \text{ we have } \mathcal{L}(u^*, p) \leq \mathcal{L}(u^*, p^*) \leq \mathcal{L}(u, p^*)$$

where

$$\mathcal{L}(u,p) := \langle p, Ku \rangle - \mathcal{J}^*(p) + \mathcal{H}(u),$$

the Lagrangian. Moreover, we can define the primal-dual gap

$$\begin{aligned} \mathcal{G}(u,p) &:= \sup_{(u',p')} \mathcal{L}(u,p') - \mathcal{L}(u',p) \\ &= \mathcal{J}(Ku) + \mathcal{H}(u) + \mathcal{J}^*(p) + \mathcal{H}^*(-K^*p), \end{aligned}$$

which vanishes iff (u,p) is a saddle point.

Example: dual ROF

$K = D$, $\mathcal{J} = \alpha\|\cdot\|_{2,1}$, $\mathcal{H} = \|\cdot - f\|_2^2/2$. Then, the dual is

$$\max_p -\mathcal{J}^*(p) - \left(\frac{1}{2}\|D^*p\|_2^2 - \langle D^*p, f\rangle\right)$$
$$= -\min_p \left(\mathcal{J}^*(p) + \frac{1}{2}\|D^*p - f\|_2^2\right) + \frac{1}{2}\|f\|^2$$

where $p \in \mathbb{R}^{m\times n\times 2}$. Here

$$\mathcal{J}^*(p) = \delta_{\{\|\cdot\|_{2,\infty}\le\alpha\}}(p) = \begin{cases} 0 & \text{if } |p_{i,j}|_2 \le \alpha \forall i,j \\ +\infty & \text{otherwise,} \end{cases}$$

and therefore the dual ROF problem is

$$\min_p\{\|D^*p - f\|_2^2 : \ |p_{i,j}|_2 \le \alpha \ \forall i,j\}.$$

From optimality conditions of saddle-point problem we have relationship between u and p:

$$u = f - D^*p.$$

Now a few algorithms . . .

Implicit gradient descent

Let J be differential. A more 'advanced' version of gradient descent is implicit gradient descent: Initial guess u^0, then iterate for $k = 0, 1, 2, \dots$

$$u^{k+1} = u^k - \tau \nabla J(u^{k+1}).$$

If u^{k+1} exists then it is a critical point of

$$\mathcal{J}(u) + \frac{\|u - u^k\|^2}{2\tau},$$

and if $\mathcal{J}$ is convex and l.s.c. then $u^{k+1} = \mathrm{prox}_{\tau \mathcal{J}}(u^k)$. If prox is easy to calculate we call $\mathcal{J}$ simple.

The prox can also make sense for non-differentiable J and the above can be generalised to subgradient descent.

Proximal point algorithm

Define Moreau-Yosida regularisation of $\mathcal{J}$ with parameter τ:

$$\mathcal{J}_\tau(\bar{u}) := \min_u \mathcal{J}(u) + \frac{\|u - \bar{u}\|^2}{2\tau}.$$

One can show

$$\nabla \mathcal{J}_\tau(\bar{u}) = \frac{\bar{u} - \mathrm{prox}_{\tau \mathcal{J}}(\bar{u})}{\tau},$$

and so, implicit gradient descent on $\mathcal{J}$

$$\begin{aligned} u^{k+1} &= \mathrm{prox}_{\tau \mathcal{J}}(u^k) \\ &= (I + \tau \partial J)^{-1}(u^k) \\ &= u^k - \tau \nabla J_\tau(u^k), \end{aligned}$$

is explicit gradient descent on $\mathcal{J}_\tau$. This is a special case of the proximal point algorithm.
Martinet 1970. Convergence rates and accelerations Bertsekas 2015; Nesterov 1983, 2004.

Forward-backward descent

Consider

$$\min_u \mathcal{J}(u) + \mathcal{H}(u),$$

with

- $\mathcal{J}$ is convex, l.s.c. and simple.
- $\mathcal{H}$ is convex with Lipschitz gradient.

Idea: Explicit descent in $\mathcal{H}$ and implicit descent in $\mathcal{J}$. That is

$$u^{k+1} = T_\tau u^k,$$

with

$$T_\tau u = \mathrm{prox}_{\tau \mathcal{J}}(u - \tau \nabla \mathcal{H}(u)).$$

Note, if u is a fixed point of T_τ then it satisfies $0 \in \nabla \mathcal{H}(u) + \partial \mathcal{J}(u)$. If $\tau \leq 1/L$ then u^k converge to a minimiser.

Accelerated version FISTA Nesterov 2004, Beck & Teboulle 2009

Primal-dual hybrid gradient

Consider

$$\min_u \mathcal{J}(Ku) + \mathcal{H}(u),$$

where $\mathcal{J}, \mathcal{H}$ are convex, l.s.c. and simple, K bounded and linear. Then, solve corresponding saddle-point problem

$$\max_p \inf_u \langle p, Ku \rangle - \mathcal{J}^*(p) + \mathcal{H}(u)$$

via

Alternate proximal descent in u and ascent in p:

$$u^{k+1} = \operatorname{prox}_{\tau\mathcal{H}}(u^k - \tau K^* p^k)$$
$$p^{k+1} = \operatorname{prox}_{\sigma\mathcal{J}^*}(p^k + \sigma K u^{k+1})$$

Arrow,Hurwicz,Uzawa 1958; Pock, Cremers, Bischof, Chambolle 2009; Esser et al. 2010
Linked to other approaches such as augmented Lagrangian and ADMM (alternating direction method of multipliers).

Primal-dual hybrid gradient

Consider

$$\min_u \mathcal{J}(Ku) + \mathcal{H}(u),$$

where $\mathcal{J}, \mathcal{H}$ are convex, l.s.c. and simple, K bounded and linear.
Then, solve corresponding saddle-point problem

$$\max_p \inf_u \langle p, Ku \rangle - \mathcal{J}^*(p) + \mathcal{H}(u)$$

via

Alternate proximal descent in u and ascent in p:

$$u^{k+1} = \mathrm{prox}_{\tau \mathcal{H}}(u^k - \tau K^* p^k)$$
$$p^{k+1} = \mathrm{prox}_{\sigma \mathcal{J}^*}(p^k + \sigma K u^{k+1})$$

Not immediately clear that this converges

Arrow,Hurwicz,Uzawa 1958; Pock, Cremers, Bischof, Chambolle 2009; Esser et al. 2010

Linked to other approaches such as augmented Lagrangian and ADMM (alternating direction method of multipliers).

Primal-dual hybrid gradient

Consider

$$\min_u \mathcal{J}(Ku) + \mathcal{H}(u),$$

where $\mathcal{J}, \mathcal{H}$ are convex, l.s.c. and simple, K bounded and linear. Then, solve corresponding saddle-point problem

$$\max_p \inf_u \langle p, Ku \rangle - \mathcal{J}^*(p) + \mathcal{H}(u)$$

via

Primal-dual hybrid gradient:

$$u^{k+1} = \text{prox}_{\tau\mathcal{H}}(u^k - \tau K^* p^k)$$
$$p^{k+1} = \text{prox}_{\sigma\mathcal{J}^*}(p^k + \sigma K(2u^{k+1} - u^k))$$

Arrow,Hurwicz,Uzawa 1958; Pock, Cremers, Bischof, Chambolle 2009; Esser et al. 2010
Linked to other approaches such as augmented Lagrangian and ADMM (alternating direction method of multipliers).

And now get your hands dirty with these approaches

...

Thank you very much for your attention!

More information see:

http://www.ccimi.maths.cam.ac.uk

http://www.cmih.maths.cam.ac.uk

http://www.damtp.cam.ac.uk/research/cia/

Email: cbs31@cam.ac.uk