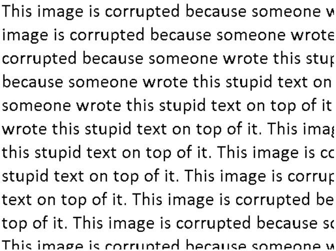
Chapter 1
Convex Analysis

Nonlinear Multiscale Methods for Image and Signal Analysis
SS 2015

Michael Moeller
Computer Vision
Department of Computer Science
TU München

Variational Problems

Example: Inpainting



$$\hat{u} = \arg\min_{u \in \mathbb{R}^n} \left\| \sqrt{(D_x u)^2 + (D_y u)^2} \right\|_1, \quad \text{such that } u_i = f_i \ \forall i \in I$$

with index set I of uncorrupted pixels.

Variational Problems

Example: Inpainting

$$\hat{u} = \arg\min_{u \in \mathbb{R}^n} \left\| \sqrt{(D_x u)^2 + (D_y u)^2} \right\|_1, \quad \text{such that } u_i = f_i \ \forall i \in I$$

with index set I of uncorrupted pixels.

Convexity

Variational Problems

Let us repeat some basics things to talk about

$$\hat{u} = \arg\min_{u \in \mathbb{R}^n} E(u).$$

Variational Problems

Let us repeat some basics things to talk about

$$\hat{u} = \arg\min_{u \in \mathbb{R}^n} E(u).$$

Definition

- For $E : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$, we call

  $$\text{dom}(E) := \{u \in \mathbb{R}^n \mid E(u) < \infty\}$$

  the domain of E.
- We call E proper if $\text{dom}(E) \neq \emptyset$.

Variational Problems

Definition: Convex Function

We call $E : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ a convex function if

1. $\text{dom}(E)$ is a convex set, i.e. for all $u, v \in \text{dom}(E)$ and all $\theta \in [0, 1]$ it holds that $\theta u + (1 - \theta)v \in \text{dom}(E)$.

2. For all $u, v \in \text{dom}(E)$ and all $\theta \in [0, 1]$ it holds that

$$E(\theta u + (1 - \theta)v) \leq \theta E(u) + (1 - \theta)E(v)$$

We call E strictly convex, if the inequality in 2 is strict for all $\theta \in]0, 1[$, and $v \neq u$.

Variational Problems

Example: Inpainting

$$\hat{u} = \arg\min_{u \in \mathbb{R}^n} \left\| \sqrt{(D_x u)^2 + (D_y u)^2} \right\|_1, \quad \text{such that } u_i = f_i \; \forall i \in I$$

with index set I of uncorrupted pixels.

$\rightarrow$ Discuss convexity.

Existence

Variational Problems

When does

$$\hat{u} = \arg\min_{u \in \mathbb{R}^n} E(u)$$

exist?

Variational Problems

When does

$$\hat{u} = \arg\min_{u \in \mathbb{R}^n} E(u)$$

exist?

- E is lower semi-continuous, i.e. for all u

  $$\liminf_{v \to u} E(v) \geq E(u)$$

  holds.
- There exists an α such that

  $$\{u \mid E(u) \leq \alpha\}$$

  is non-empty and bounded.

Proof: Board.

Variational Problems

Fundamental Theorem of Optimization

If $E : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is lower semi-continuous and has a nonempty bounded sublevelset, then there exists

$$\hat{u} = \arg\min_{u \in \mathbb{R}^n} E(u)$$

Variational Problems

Fundamental Theorem of Optimization

If $E : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is lower semi-continuous and has a nonempty bounded sublevelset, then there exists

$$\hat{u} = \arg\min_{u \in \mathbb{R}^n} E(u)$$

Remark: For a proper convex function, lower semi-continuity is the same as the closedness of the sublevelsets.

Variational Problems

Fundamental Theorem of Optimization

If $E : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is lower semi-continuous and has a nonempty bounded sublevelset, then there exists

$$\hat{u} = \arg\min_{u \in \mathbb{R}^n} E(u)$$

Remark: For a proper convex function, lower semi-continuity is the same as the closedness of the sublevelsets.

Examples on the board:

- A convex continuous function that does not have a minimizer
- A convex function with bounded sublevelsets that does not have a minimizer

Variational Problems

Continuity of Convex Functions

If $E : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is convex, then E is locally Lipschitz (and hence continuous) on int(dom(E)).

Proof: Exercise (in 1d)

Board: Considering the interior is important!

Variational Problems

Continuity of Convex Functions

If $E : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is convex, then E is locally Lipschitz (and hence continuous) on int(dom(E)).

Proof: Exercise (in 1d)

Board: Considering the interior is important!

Conclusion

If $E : \mathbb{R}^n \to \mathbb{R}$ is convex, then E is continuous.

Variational Problems

Definition

We call $E : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ coercive, if all sequences $(u_n)_n$ with $\|u_n\| \to \infty$ meet $E(u_n) \to \infty$.

Theorem

If $E : \mathbb{R}^n \to \mathbb{R}$ is convex and coercive, then there exists

$$\hat{u} = \arg\min_{u \in \mathbb{R}^n} E(u).$$

Variational Problems

When is

$$\hat{u} = \arg\min_{u \in \mathbb{R}^n} E(u)$$

unique?

Variational Problems

When is

$$\hat{u} = \arg\min_{u \in \mathbb{R}^n} E(u)$$

unique?

Theorem

If $E : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is convex, then any local minimum is a global minimum. If E is strictly convex, the global minimum is unique.

Subdifferential Calculus

Variational Problems

What is an optimality condition for

$$\hat{u} = \arg\min_{u \in \mathbb{R}^n} E(u)?$$

Variational Problems

What is an optimality condition for

$$\hat{u} = \arg\min_{u \in \mathbb{R}^n} E(u)?$$

Definition: Subdifferential

We call

$$\partial E(u) = \{p \in \mathbb{R}^n \mid E(v) - E(u) - \langle p, v - u \rangle \geq 0\}$$

the subdifferential of E at u.

- Elements of $\partial E(u)$ are called subgradients.
- If $\partial E(u) \neq \emptyset$, we call E subdifferentiable at E.
- By convention, $\partial E(u) = \emptyset$ for $u \neq \mathrm{dom}(E)$.

Variational Problems

What is an optimality condition for

$$\hat{u} = \arg\min_{u \in \mathbb{R}^n} E(u)?$$

Definition: Subdifferential

We call

$$\partial E(u) = \{p \in \mathbb{R}^n \mid E(v) - E(u) - \langle p, v - u \rangle \geq 0\}$$

the subdifferential of E at u.

- Elements of $\partial E(u)$ are called subgradients.
- If $\partial E(u) \neq \emptyset$, we call E subdifferentiable at E.
- By convention, $\partial E(u) = \emptyset$ for $u \neq \mathrm{dom}(E)$.

Theorem: Optimality condition

Let $0 \in \partial E(\hat{u})$. Then $\hat{u} \in \arg\min_u E(u)$.

Variational Problems

Examples for non-differentiable functions:

- The ℓ^1 norm.

Variational Problems

Examples for non-differentiable functions:

- The ℓ^1 norm.
- Functional

$$E(u) = \begin{cases} 0 & \text{if } u \geq 0 \\ \infty & \text{else.} \end{cases}$$

Variational Problems

Examples for non-differentiable functions:

- The ℓ^1 norm.
- Functional

$$E(u) = \begin{cases} 0 & \text{if } u \geq 0 \\ \infty & \text{else.} \end{cases}$$

Subdifferential and derivatives

Let the convex function $E : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be differentiable at $x \in \mathrm{dom}(E)$. Then

$$\partial E(x) = \{\nabla E(x)\}.$$

Variational Problems

Is any convex E subdifferentiable at $x \in \text{dom}(E)$?

[1] Rockafellar, Convex Analysis, Theorem 23.4

Variational Problems

Is any convex E subdifferentiable at $x \in \text{dom}(E)$?

Answer: Almost...

Definition: Relative Interior

The *relative interior* of a convex set M is defined as

$$\text{ri}(M) := \{x \in M \mid \forall y \in M,\ \exists \lambda > 1,\ \text{s.t. } \lambda x + (1-\lambda) y \in M\}$$

[1] Rockafellar, Convex Analysis, Theorem 23.4

Variational Problems

Is any convex E subdifferentiable at $x \in \mathrm{dom}(E)$?

Answer: Almost...

Definition: Relative Interior

The *relative interior* of a convex set M is defined as

$$\mathrm{ri}(M) := \{x \in M \mid \forall y \in M,\ \exists \lambda > 1,\ \text{s.t. } \lambda x + (1-\lambda) y \in M\}$$

Theorem: Subdifferentiability[1]

If E is a proper convex function and $u \in \mathrm{ri}(\mathrm{dom}(E))$, then $\partial E(u)$ is non-empty and bounded.

[1] Rockafellar, Convex Analysis, Theorem 23.4

Variational Problems

Theorem: Sum rule[2]

Let E_1, E_2 be convex functions such that

$$\text{ri}(\text{dom}(E_1)) \cap \text{ri}(\text{dom}(E_2)) \neq \emptyset,$$

then it holds that

$$\partial(E_1 + E_2)(u) = \partial E_1(u) + \partial E_2(u).$$

[2] Rockafellar, Convex Analysis, Theorem 23.8

Variational Problems

Theorem: Sum rule[2]

Let E_1, E_2 be convex functions such that

$$\mathrm{ri}(\mathrm{dom}(E_1)) \cap \mathrm{ri}(\mathrm{dom}(E_2)) \neq \emptyset,$$

then it holds that

$$\partial(E_1 + E_2)(u) = \partial E_1(u) + \partial E_2(u).$$

Example: Minimize $(u - f)^2 + \iota_{u \geq 0}(u)$.

[2] Rockafellar, Convex Analysis, Theorem 23.8

Variational Problems

Theorem: Sum rule[2]

Let E_1, E_2 be convex functions such that

$$\mathrm{ri}(\mathrm{dom}(E_1)) \cap \mathrm{ri}(\mathrm{dom}(E_2)) \neq \emptyset,$$

then it holds that

$$\partial(E_1 + E_2)(u) = \partial E_1(u) + \partial E_2(u).$$

Example: Minimize $(u - f)^2 + \iota_{u \geq 0}(u)$.

Example: Minimize $0.5(u - f)^2 + \alpha|u|$.

[2] Rockafellar, Convex Analysis, Theorem 23.8

Variational Problems

Theorem: Chain rule[3]

If $A \in \mathbb{R}^{m \times n}$, $E : \mathbb{R}^m \to \mathbb{R} \cup \{\infty\}$ is convex, and $\text{ri}(\text{dom}(E)) \cap \text{range}(A) \neq \emptyset$, then

$$\partial(E \circ A)(u) = A^* \partial E(Au)$$

---

[3] Rockafellar, Convex Analysis, Theorem 23.9

Variational Problems

Theorem: Chain rule[3]

If $A \in \mathbb{R}^{m\times n}$, $E : \mathbb{R}^m \to \mathbb{R} \cup \{\infty\}$ is convex, and $\text{ri}(\text{dom}(E)) \cap \text{range}(A) \neq \emptyset$, then

$$\partial(E \circ A)(u) = A^* \partial E(Au)$$

Example: Minimize $\|Au - f\|_2^2$.

[3] Rockafellar, Convex Analysis, Theorem 23.9

Variational Problems

Summary (without assumptions):

- $\partial E(u) = \{p \in \mathbb{R}^n \mid E(v) - E(u) - \langle p, v - u \rangle \geq 0\}$
- If E differentiable: $\partial E(x) = \{\nabla E(x)\}$
- Sum rule $\partial(E_1 + E_2)(x) = \partial E_1(x) + \partial E_2(x)$
- Cain rule $\partial(E \circ A)(u) = A^* \partial E(Au)$

TV minimization

TV minimization

TV minimization

What is TV again?

For $u \in \mathbb{R}^{m \times n}$ let us consider the anisotropic total variation

$$TV_a(u) = \sum_{i=2}^{m} \sum_{j=2}^{n} |u_{i,j} - u_{i-1,j}| + |u_{i,j} - u_{i,j-1}|$$

What is TV again?

For $u \in \mathbb{R}^{m \times n}$ let us consider the anisotropic total variation

$$TV_a(u) = \sum_{i=2}^{m} \sum_{j=2}^{n} |u_{i,j} - u_{i-1,j}| + |u_{i,j} - u_{i,j-1}|$$

For doing math, it is often easier to consider $\vec{u}_{i+m(j-1)} = u(i,j)$ and write

$$TV_a(u) = \|K\vec{u}\|_1$$

for a suitable matrix K that discretizes the gradient.

TV Minimization

Our problem becomes

$$u(\alpha) = \arg \min_{u \in \mathbb{R}^{nm}} \frac{1}{2}\|u - f\|_2^2 + \alpha \|Ku\|_1.$$

Let us try to apply all the learned theory. The minimizer is obtained at

$$0 \in u(\alpha) - f + \alpha K^T q$$

with $q \in \partial \|Ku(\alpha)\|_1$, i.e.

$$q_i \begin{cases} = 1 & \text{if } (Ku(\alpha))_i > 0 \\ = -1 & \text{if } (Ku(\alpha))_i < 0 \\ \in [-1, 1] & \text{if } (Ku(\alpha))_i = 0 \end{cases}$$

Seems extremely difficult to find...

TV Minimization

Crazy idea:

$$\min_u \frac{1}{2}\|u - f\|_2^2 + \alpha\|Ku\|_1 = \min_u \frac{1}{2}\|u - f\|_2^2 + \alpha \sup_{\|q\|_\infty \leq 1} \langle Ku, q\rangle$$

$$= \min_u \sup_{\|q\|_\infty \leq 1} \frac{1}{2}\|u - f\|_2^2 + \alpha\langle Ku, q\rangle$$

Can we exchange min and sup?

TV Minimization

Saddle point problems[4]

Let C and D be non-empty closed convex sets in $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively, and let S be a continuous finite concave-convex function on $C \times D$. If either C or D is bounded, one has

$$\inf_{v \in D} \sup_{q \in C} S(v, q) = \sup_{q \in C} \inf_{v \in D} S(v, q).$$

[4] Rockafellar, Convex Analysis, Corollary 37.3.2

TV Minimization

Saddle point problems[4]

Let C and D be non-empty closed convex sets in $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively, and let S be a continuous finite concave-convex function on $C \times D$. If either C or D is bounded, one has

$$\inf_{v \in D} \sup_{q \in C} S(v,q) = \sup_{q \in C} \inf_{v \in D} S(v,q).$$

We can therefore compute

$$\begin{aligned}\min_u \frac{1}{2}\|u-f\|_2^2 + \alpha\|Ku\|_1 &= \min_u \sup_{\|q\|_\infty \leq 1} \frac{1}{2}\|u-f\|_2^2 + \alpha\langle Ku, q\rangle \\ &= \sup_{\|q\|_\infty \leq 1} \min_u \frac{1}{2}\|u-f\|_2^2 + \alpha\langle Ku, q\rangle.\end{aligned}$$

[4] Rockafellar, Convex Analysis, Corollary 37.3.2

TV Minimization

Now the inner minimization problem obtains its optimum at

$$\begin{aligned} 0 &= u - f + \alpha K^T q, \\ \Rightarrow u &= f - \alpha K^T q. \end{aligned}$$

The remaining problem in q becomes

$$\begin{aligned} &\sup_{\|q\|_\infty \leq 1} \frac{1}{2}\|f - \alpha K^T q - f\|_2^2 + \alpha \langle K(f - \alpha K^T q), q \rangle \\ = &\sup_{\|q\|_\infty \leq 1} \frac{1}{2}\|\alpha K^T q\|_2^2 + \alpha \langle K f, q \rangle - \|\alpha K^T q\|_2^2 \\ = &\sup_{\|q\|_\infty \leq 1} -\frac{1}{2}\|\alpha K^T q - f\|_2^2 \end{aligned}$$

TV Minimization

Since we prefer minimizations over maximizations, we write

$$\hat{q} = \arg \max_{\|q\|_\infty \leq 1} -\frac{1}{2}\|\alpha K^T q - f\|_2^2$$
$$= \arg \min_{\|q\|_\infty \leq 1} \frac{1}{2}\left\|K^T q - \frac{f}{\alpha}\right\|_2^2$$

Idea: Gradient descent + project onto feasible set.

$$q^{k+1} = \pi_{\|q\|_\infty \leq 1}\left(q^k - \tau K\left(K^T q^k - \frac{f}{\alpha}\right)\right)$$

TV Minimization

Gradient projection algorithm[5]

The algorithm

$$q^{k+1} = \pi_{\|\cdot\|_\infty \leq 1}\left(q^k - \tau K\left(K^T q^k - \frac{f}{\alpha}\right)\right)$$

with $u^k = f - \alpha q^k$, for TV minimization converges for $\tau < \frac{1}{4}$.

Remark: The $1/4$ is two over the Lipschitz constant of the gradient of the smooth objective.

[5] Levitin, Polyak, *Constrained minimization problems*, 1966. Goldstein, *Convex programming in Hilbert space*, 1964.