Chapter 1

Convex Analysis

*Nonlinear Multiscale Methods for Image and Signal Analysis*

SS 2015

Michael Moeller

Computer Vision

Department of Computer Science

TU München

updated 07.05.2015
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

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

with index set $I$ of uncorrupted pixels.
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
  \[ \text{dom}(E) := \{ u \in \mathbb{R}^n | E(u) < \infty \} \]
  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

This image is corrupted because someone wrote this stupid text on top of it. This image is corrupted because someone wrote this stupid text on top of it. This image is corrupted because someone wrote this stupid text on top of it. This image is corrupted because someone wrote this stupid text on top of it. This image is corrupted because someone wrote this stupid text on top of it. This image is corrupted because someone wrote this stupid text on top of it. This image is corrupted because someone wrote this stupid text on top of it.

This image is corrupted because someone wrote this stupid text on top of it. This image is corrupted because someone wrote this stupid text on top of it. This image is corrupted because someone wrote this stupid text on top of it. This image is corrupted because someone wrote this stupid text on top of it. This image is corrupted because someone wrote this stupid text on top of it. This image is corrupted because someone wrote this stupid text on top of it. This image is corrupted because someone wrote this stupid text on top of it.

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

with index set $I$ of uncorrupted pixels.

→ Discuss convexity.
Existence
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 $\alpha$ such that

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

  is non-empty and bounded.

Proof: Board.
**Variational Problems**

**Fundamental Theorem of Optimization**

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

Proof: Exercise (in 1d)

Board: Considering the interior is important!

Conclusion

If \( E : \mathbb{R}^n \rightarrow \mathbb{R} \) is convex, then \( E \) is continuous.
**Variational Problems**

**Definition**

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

**Theorem**

If $E : \mathbb{R}^n \rightarrow \mathbb{R}$ 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?

**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

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

**Definition: Subdifferential**

We call

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

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 $u \neq \text{dom}(E)$.

**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 $\ell^1$ 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 $x \in \text{dom}(E)$. Then

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

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

**Theorem: Subdifferentiability**

If $E$ is a proper convex function and $u \in \text{ri}(	ext{dom}(E))$, 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

\[
\partial(E_1 + E_2)(u) = \partial E_1(u) + \partial E_2(u).
\]

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**

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 $\|Au - f\|_2^2$.

---

*Rockafellar, Convex Analysis, Theorem 23.9*
Variational Problems

Summary (without assumptions):

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

• 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

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

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

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

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?
Saddle point problems

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

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

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

---

4 Rockafellar, Convex Analysis, Corollary 37.3.2
TV Minimization

Now the inner minimization problem obtains its optimum at

\[ 0 = u - f + \alpha K^T q, \]
\[ \Rightarrow u = f - \alpha K^T q. \]

The remaining problem in \( q \) becomes

\[
\sup_{\|q\|_\infty \leq 1} \frac{1}{2} \|f - \alpha K^T q - f\|^2 + \alpha \langle K(f - \alpha K^T q), q \rangle
= \sup_{\|q\|_\infty \leq 1} \frac{1}{2} \|\alpha K^T q\|^2 + \alpha \langle Kf, q \rangle - \|\alpha K^T q\|^2
= \sup_{\|q\|_\infty \leq 1} -\frac{1}{2} \|\alpha K^T q - f\|^2
\]
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.

---

TV Minimization

```matlab
1  img = im2double(imread('cameraman.tif'));
2  [m,n] = size(img);
3
4  e = ones(m,1);
5  yDerMat = spdiags([-e e], 0:1, m, m);
6  yDerMat(end) =0;
7
8  e = ones(n,1);
9  xDerMat = spdiags([-e e], 0:1, n, n)';
10  xDerMat(end)=0;
11
12  yDer = yDerMat*img;
13  xDer = img*xDerMat;
14
15  gradientMatrixForVectorizedImage = ...
16      [kron(xDerMat',speye(m,m)); kron(speye(m,m), ...
17          yDerMat)];
18
19  imgGradient = gradientMatrixForVectorizedImage*img(:);
20  figure, imagesc(reshape(imgGradient(1:n*m), ...
21      [n,m])), colorbar
```

Note that $AXB = C \iff \text{kron}(B', A)\vec{X} = \vec{C}$!
Duality
We replaced $\|Ku\|_1$ (not differentiable) with $\sup_{\|q\|_\infty \leq 1} \langle q, Ku \rangle$.

Numerics became easier.

Is there a systematic concept behind this idea?
Variational Problems

Very important concept: Duality!

**Definition: Convex Conjugate**

We define the convex conjugate of the function $E : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ to be

$$E^*(p) = \sup_{u \in \mathbb{R}^n} (\langle u, p \rangle - E(u)).$$

Board: $E^*$ is convex.
Variational Problems

Examples:

• $E(u) = u^2$ leads to $E^*(p) = p^2$

• $E(u) = \|u\|_2$ leads to $E^*(p) = \begin{cases} 0 & \text{if } \|p\|_2 \leq 1, \\ \infty & \text{else.} \end{cases}$

• $E(u) = \|u\|_1$ leads to $E^*(p) = \begin{cases} 0 & \text{if } \|p\|_\infty \leq 1, \\ \infty & \text{else.} \end{cases}$

• $E(u) = \|u\|_\infty$ leads to $E^*(p) = \begin{cases} 0 & \text{if } \|p\|_1 \leq 1, \\ \infty & \text{else.} \end{cases}$

• $E(u) = \begin{cases} 0 & \text{if } \|u\|_2 \leq 1, \\ \infty & \text{else.} \end{cases}$ leads to $E^*(p) = \|p\|_2$.

• $E(u) = \begin{cases} 0 & \text{if } \|u\|_\infty \leq 1, \\ \infty & \text{else.} \end{cases}$ leads to $E^*(p) = \|p\|_1$.

• $E(u) = \begin{cases} 0 & \text{if } \|u\|_1 \leq 1, \\ \infty & \text{else.} \end{cases}$ leads to $E^*(p) = \|p\|_\infty$.

Suspicion: $E^{**} = E$?
**Variational Problems**

**Fenchel-Young Inequality**\(^6\)

Let \( E \) be proper, convex and lower semi-continuous, \( u \in \text{dom}(E) \subset \mathbb{R}^n \), and \( p \in \mathbb{R}^n \), then

\[
E(u) + E^*(p) \geq \langle u, p \rangle.
\]

Equality holds if and only if \( p \in \partial E(u) \).

**Theorem: Biconjugate**\(^7\)

Let \( E \) be proper, convex and lower semi-continuous, then

\( E^{**} = E \).

Now we understand what we did for TV minimization...

---

\(^6\)Borwein, Zhu *Techniques of variational analysis*, Proposition 4.4.1

\(^7\)Rockafellar, Convex Analysis, Theorem 12.2
Theorem: Subgradient of convex conjugate\(^8\)

Let \( E \) be proper, convex and lower semi-continuous, then the following two conditions are equivalent:

- \( p \in \partial E(u) \)
- \( u \in \partial E^*(p) \)

Board: Example with proximity operator.

---

\(^8\)Rockafellar, Convex Analysis, Theorem 23.5
Variational Problems

What is this good for? Easier analysis and optimization!

Fenchel’s Duality Theorem\(^9\)

The following problems yield the same result:

\[
\begin{align*}
\inf_u & \quad H(u) + R(Ku) \quad & \text{"Primal"} \\
\inf_u \sup_q & \quad H(u) + \langle q, Ku \rangle - R^*(q) \quad & \text{"Saddle point"} \\
\sup_q \inf_u & \quad H(u) + \langle q, Ku \rangle - R^*(q) \\
\sup_q & \quad -H^*(-K^*q) - R^*(q) \quad & \text{"Dual"}
\end{align*}
\]

Assuming that \(H\) and \(R\) are proper, lower semi-continuous, convex functions with \(\text{ri}(\text{dom}(H)) \cap \text{ri}(\text{dom}(R \circ K)) \neq \emptyset\).

\(^9\)C.f. Rockafellar, Convex Analysis, Section 31, or Borwein, Zhu Techniques of variational analysis, Theorem 4.4.3
Variational Problems
Consider the isotropic total variation

\[ TV(u) = \sum_{i=2}^{m} \sum_{j=2}^{n} \sqrt{(u_{i,j} - u_{i-1,j})^2 + (u_{i,j} - u_{i,j-1})^2} = \|Ku\|_{2,1} \]

How do we minimize

\[ \frac{1}{2} \| u - f \|_2^2 + \alpha \|Ku\|_{2,1} ? \]

Dual problem:

\[ \frac{1}{2} \left\| K^T q - \frac{f}{\alpha} \right\|_2^2 + (\| \cdot \|_{2,1})^*(q) \]

Similar to previous examples:

\[ (\| \cdot \|_{2,1})^*(q) = \begin{cases} 0 & \text{if } \|q\|_{2,\infty} \leq 1 \ \forall i, \\ \infty & \text{else.} \end{cases} \]

→ Gradient projection algorithm!