Total variation denoising with iterated conditional expectation

Cécile Louchet[1] and Lionel Moisan[2]

[1] Université d'Orléans, Institut Denis Poisson, France
[2] Université Paris Descartes, MAP5, France

Workshop IHP "Statistical modeling for shapes and imaging"
March, 12th 2019

TV restoration of images

Image formation model

$$v = A\mathbf{u} + n$$

- $v \in \mathbb{R}^{\Omega'}$: observed image
- $A : \mathbb{R}^{\Omega} \to \mathbb{R}^{\Omega'}$: linear operator
  ($A = Id \to$ denoising; $A = k * \cdot \to$ deblurring...)
- n: Gaussian additive white noise $\sim \mathcal{N}(0, \sigma^2)$
- $\mathbf{u} \in \mathbb{R}^{\Omega}$: image that we want to estimate.

Rudin-Osher-Fatemi image recovery

Choose $\quad \hat{u}_{\text{ROF}} = \arg\min_{u \in \mathbb{R}^{\Omega}} \mathcal{E}(u) := \|Au - v\|^2 + \lambda TV(u)$

- Total Variation: $TV(u) = \|\nabla u\|_1$
- $\lambda \geq 0$ is a user-controlled regularity parameter.

TV restoration of images

Bayesian viewpoint

$\hat{u}_{\text{ROF}}$ is a Maximum A Posteriori in a Bayes framework:

$$\begin{aligned}
\hat{u}_{\text{ROF}} &= \arg\min_u \|Au - v\|^2 + \lambda TV(u) \\
&= \arg\max_u \frac{1}{Z} e^{-\frac{\|Au-v\|^2}{2\sigma^2}} e^{-\beta TV(u)} \qquad (\text{where } \beta = \frac{\lambda}{2\sigma^2}) \\
&= \arg\max_u P(v|u)\, P(u) = \arg\max_u P(u|v)
\end{aligned}$$

$$\text{with } \begin{cases} P(v|u) = \frac{1}{Z} e^{-\frac{\|Au-v\|^2}{2\sigma^2}} & \text{(image formation model)} \\ P(u) = \frac{1}{Z'} e^{-\beta TV(u)} & \text{(prior distribution)} \end{cases}$$

Restoration with TV-LSE

We have $\hat{u}_{\mathrm{ROF}} = \arg\max_u P(u|v)$.

Definition: image restoration by TV-**L**east **S**quare **E**stimator [1]

$$\hat{u}_{\mathrm{LSE}} = \mathbb{E}[u|v] = \frac{1}{Z}\int_{\mathbb{R}^\Omega} u\, e^{-\frac{1}{2\sigma^2}(\|Au-v\|^2+\lambda TV(u))}\, du$$

No staircasing in LSE denoising ($A = Id$)

$\forall x, y \in \Omega$, the set $\{v \in \mathbb{R}^\Omega : \hat{u}_{\mathrm{LSE}}(x) = \hat{u}_{\mathrm{LSE}}(y)\}$ has measure 0.

Computation of TV-LSE

For each $x \in \Omega$, $\quad \hat{u}_{\mathrm{LSE}}(x) = \frac{1}{Z}\int_{\mathbb{R}^\Omega} u(x)\, e^{-\frac{1}{2\sigma^2}\mathcal{E}(u)}\, du.$

- integral on $\mathbb{R}^\Omega$ where $|\Omega|$ = number of pixels $\approx 10^6$...
- MCMC techniques but with convergence in $O(1/\sqrt{N})$.

[1] L., Moisan (2013). Posterior expectation of the total variation model: Properties and experiments. *SIIMS*.

Outline

1. TV denoising with Iterated Conditional Expectations

2. Other (imaging?) tasks with ICE
   - Deblurring and inverse problems regularized with TV
   - TV-ICE denoising for Poisson noise
   - ICE of a convex functional
   - ICE of a convex set

Outline

1. TV denoising with Iterated Conditional Expectations

2. Other (imaging?) tasks with ICE
   - Deblurring and inverse problems regularized with TV
   - TV-ICE denoising for Poisson noise
   - ICE of a convex functional
   - ICE of a convex set

The idea of TV-ICE denoising

Recall in the case $A = Id$:

$$\hat{u}_{\mathrm{LSE}}(x) = \frac{\int_{\mathbb{R}^\Omega} u(x)\, e^{-\frac{\|u-v\|^2+\lambda TV(u)}{2\sigma^2}}\, du}{\int_{\mathbb{R}^\Omega} e^{-\frac{\|u-v\|^2+\lambda TV(u)}{2\sigma^2}}\, du}$$

Idea: integrating one variable at a time

$$\frac{\int_{\mathbb{R}} u(x)\, e^{-\frac{\|u-v\|^2+\lambda TV(u)}{2\sigma^2}}\, du(x)}{\int_{\mathbb{R}} e^{-\frac{\|u-v\|^2+\lambda TV(u)}{2\sigma^2}}\, du(x)} =: f_{v(x)}(u(\mathcal{N}_x))$$

- This is the posterior expectation of $u(x)$ conditionally to $u(x^c)$.
- It depends on the values of $u(x^c)$. But we can iterate: convergence hopefully?
- From now on: $\mathcal{N}_x$ = 4-neighbor system and

$$TV(u) = \frac{1}{2} \sum_{x\in\Omega} \sum_{y\in\mathcal{N}_x} |u(y) - u(x)|.$$

One iteration: closed formula

If $u(\mathcal{N}_x) = \{a, b, c, d\}$ with $a \leq b \leq c \leq d$ and if $v(x) = t$, then
For any $n \in \mathbb{N}^*$, for any sorted n-uple (a_j) and for any n-uple (β_j), we have

$$f_t(u(\mathcal{N}_x)) = t - \frac{\sum_{i=0}^{n} \mu_i I^t_{\mu_i,\nu_i}(a_i, a_{i+1})}{\sum_{i=0}^{n} I^t_{\mu_i,\nu_i}(a_i, a_{i+1})}$$

where $\{a_1, \ldots, a_4\} = \{u(\mathcal{N}_x)\}$ and
$-\infty = a_0 \leq a_1 \leq \cdots \leq a_5 = +\infty$,

$$\forall i, \quad \mu_i = \frac{\lambda}{2} \sum_{j=1}^{n} \varepsilon_{i,j}, \quad \nu_i = -\lambda \sum_{j=1}^{n} \varepsilon_{i,j} a_j \quad \varepsilon_{i,j} = \begin{cases} 1 & \text{if } i \geq j \\ -1 & \text{otherwise} \end{cases}$$

and

$$I^t_{\mu,\nu}(a, b) = \left(\operatorname{erf}\left(\frac{b - t + \mu}{\sigma\sqrt{2}} \right) - \operatorname{erf}\left(\frac{a - t + \mu}{\sigma\sqrt{2}} \right) \right) e^{-\frac{1}{2\sigma^2}(2\mu t - \mu^2 + \nu)}.$$

Theorem and definition of TV-ICE

Consider an image $v : \Omega \to \mathbb{R}$ and $\lambda, \sigma > 0$.
The sequence $(u^n)_{n \geq 0}$ defined recursively by u^0 and

$$\forall x \in \Omega, \quad u^{n+1}(x) = f_{v(x)}(u^n(\mathcal{N}_x))$$

converges *linearly* to an image $\hat{u}_{\mathrm{ICE}}$ independent of u^0 and satisfies

$$\forall x \in \Omega, \quad \hat{u}_{\mathrm{ICE}}(x) = \mathbb{E}_{u|v}[u(x) \mid u(x^c) = \hat{u}_{\mathrm{ICE}}(x^c)].$$

Idea of the proof: we define F_v by $u^{n+1} = F_v(u^n)$. Then

$$F_v(u)(x) = f_{v(x)}(u(\mathcal{N}_x)).$$

- F_v is $\mathcal{C}^1$ and monotone: $w_1 \leq w_2 \Rightarrow F_v(w_1) \leq F_v(w_2)$
- $f_{t-c}(w(\mathcal{N}_x) - c) = f_t(w(\mathcal{N}_x)) - c$ and implies $\|\mathrm{Jac}\ F_v\|_\infty < 1$
- $K_w = \left[\min(\min_\Omega v, \min_\Omega w), \max(\max_\Omega v, \max_\Omega w)\right]^\Omega$ satisfies $F_v(K_w) \subset K_w$ for any w.

Properties of TV-ICE denoising

ICE is not LSE.

Proof: LSE is a prox, ICE is not.

No staircasing

Let x and y be neighbor pixels. The set $\{v \in \mathbb{R}^\Omega \ : \ \hat{u}_{\mathrm{ICE}}(x) = \hat{u}_{\mathrm{ICE}}(y)\}$ has measure 0.

Proof: $v \mapsto \hat{u}_{\mathrm{ICE}}$ is $\mathcal{C}^1$.

Recovery of edges

$$v(x) - 2\lambda \leq \hat{u}_{\mathrm{ICE}}(x) \leq v(x) + 2\lambda.$$

$\rightarrow$ a strong local contrast essentially persists.

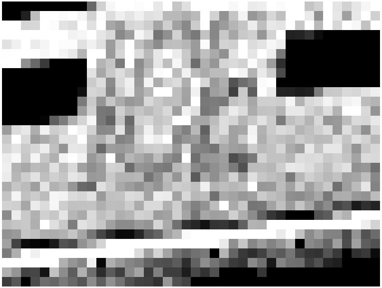
Proof: $f_t(a, b, c, d)$ is a weighted average (with positive coefficients) of t, $t \pm \lambda$, and $t \pm 2\lambda$, it belongs to $[t - 2\lambda, t + 2\lambda]$. This latter property is shared with $\hat{u}_{\mathrm{ROF}}$ and $\hat{u}_{\mathrm{LSE}}$.

noisy

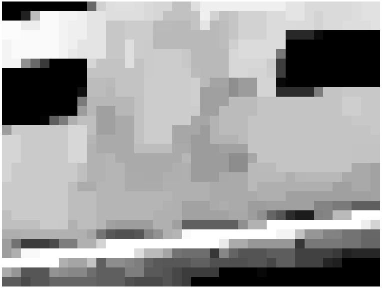
ROF

ICE

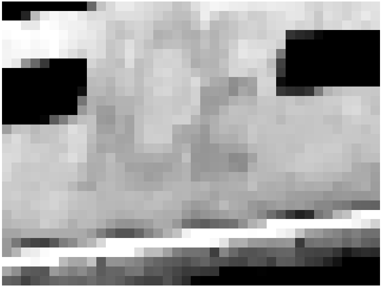
LSE



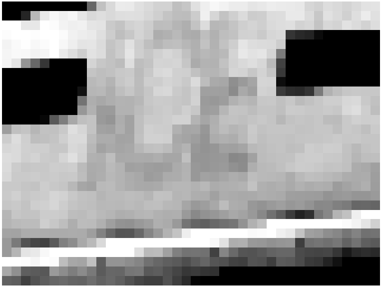
noisy



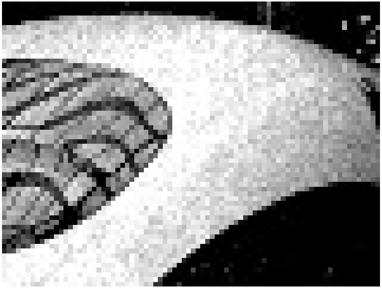
ROF



ICE



LSE

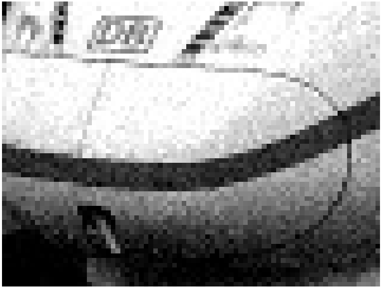


noisy

ROF

ICE

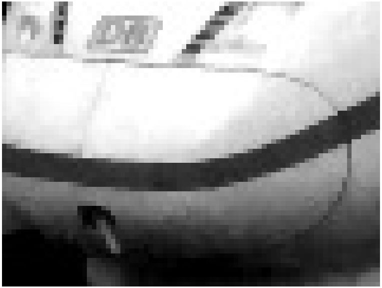
LSE



noisy

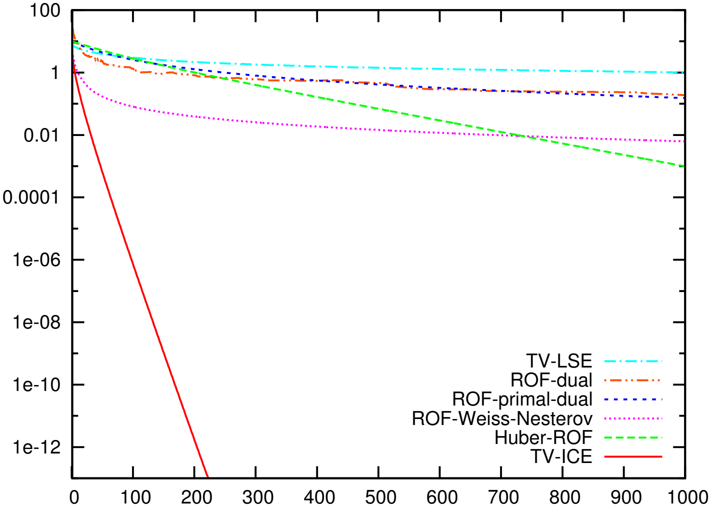
ROF

ICE



LSE

Convergence curves for different algorithms of TV-denoising



Outline

1. TV denoising with Iterated Conditional Expectations

2. Other (imaging?) tasks with ICE
   - Deblurring and inverse problems regularized with TV
   - TV-ICE denoising for Poisson noise
   - ICE of a convex functional
   - ICE of a convex set

Outline

1. TV denoising with Iterated Conditional Expectations

2. Other (imaging?) tasks with ICE
   - Deblurring and inverse problems regularized with TV
   - TV-ICE denoising for Poisson noise
   - ICE of a convex functional
   - ICE of a convex set

The idea of TV-ICE restoration

Definition of the ICE sequence

Start with an arbitrary u^0 and for all $n \in \mathbb{N}$ set

$$u^{n+1}(x) = \frac{1}{Z} \int_{\mathbb{R}} u^n(x) e^{-\frac{\|Au^n - v\|^2 + \lambda TV(u^n)}{2\sigma^2}} \, du^n(x).$$

- computable?
- convergence $u^n \to \hat{u}_{\mathrm{ICE}}$?

The iterations are easy to deduce from the denoising case!

Case where $\|A\delta_x\|^2$ does not depend on x : we have

$$\begin{cases} w^{n+1} = u^n - \gamma A^*(Au^n - v) \\ u^{n+1} = F_{w^{n+1}}(u^n) \end{cases}$$

with parameters $(\gamma\sigma^2, \gamma\lambda)$, where $\gamma = \|A\delta_x\|^{-2}$.

Convergence condition

Assumptions

- $\gamma = \|A\delta_x\|^{-2}$ does not depend on x
- $A\mathbb{1}_\Omega = \mathbb{1}_{\Omega'}$

Theorem

If $\gamma < 2$, then $(u^n)_{n\in\mathbb{N}}$ linearly converges to a limit $\hat{u}_{\mathrm{ICE}}$ independent of u^0 such that

$$\forall x \in \Omega, \quad \hat{u}_{\mathrm{ICE}}(x) = \mathbb{E}_{u|v}[u(x) \,|\, u(x^c) = \hat{u}_{\mathrm{ICE}}(x^c)].$$

But for each $\gamma \geq 2$ there are always cases of non-convergence.

- deconvolution: if $A = k * \cdot$ with $\sum k = 1$, then $\gamma = 1/\|k\|^2$.
  Gaussian blur: $\gamma < 2 \Leftrightarrow \sigma_A \lessapprox 0.5$ pixel
  $\rightarrow$ k should be very concentrated.
- zooming: if $A =$ block-averaging, blocks should have size < 2.

$\rightarrow$ very limited applications!

4 possible strategies to ensure convergence

$$\begin{cases} w^{n+1} = u^n - \gamma A^*(Au^n - v) = (I - \gamma A^* A)u^n + \gamma A^* v & (1) \\ u^{n+1}(x) = F_{w^{n+1}}(u^n) & (2) \end{cases}$$

1st strategy: averaging on u.

Replace (2) step with $u^{n+1} = (1-r)u^n + r\,F_{w^{n+1}}(u^n)$
Observation: $r \leq \min(1, \frac{2}{\gamma\rho(A^*A)})$ $\implies$ linear convergence.

2nd strategy: averaging on w.

Replace (1) step with $w^{n+1} = (1-s)w^n + s\,(u^n - \gamma A^*(Au^n - v))$
Observation: $s \leq \min(1, \frac{2}{\gamma\rho(A^*A)})$ $\implies$ linear convergence.

3rd strategy: set γ free.

Replace (1) step with $w^{n+1} = u^n - \tau A^*(Au^n - v)$, $\tau > 0$.
Observation: $\tau < 2 \implies$ linear convergence.

4th strategy: "implicitize".

Replace (1) step with with $w^{n+1} = (I + \gamma A^* A)^{-1}(u^n + \gamma A^* v)$.
Observation: linear convergence.

Application to image deblurring

Framework

$$A = k * \cdot \quad \Rightarrow \quad \gamma = 1/\|k\|^2$$

- If $\|k\|^2 > 1/2$, the natural strategy applies

$$\begin{cases} w^{n+1} = u^n - \gamma \check{k} * (k * u^n - v) \\ u^{n+1} = F_{w^{n+1}}(u^n) \end{cases}$$

- else, the averaging and free-gamma strategies always apply

$$\begin{cases} w^{n+1} = (1-s)w^n + s\left(u^n - \tau \check{k} * (k * u^n - v)\right) \\ u^{n+1} = (1-r)u^n + r\, F_{w^{n+1}}(u^n) \end{cases}$$

- Implicit strategy available only for periodic boundary conds:

$$\begin{cases} w^{n+1} = (I + \gamma A^* A)^{-1}(u^n + \gamma A^* v) = \mathcal{F}^{-1}\left(\frac{\mathcal{F}u + \gamma(\mathcal{F}k)^* \cdot \mathcal{F}(v)}{1 + \gamma|\mathcal{F}k|^2}\right) \\ u^{n+1} = F_{w^{n+1}}(u^n). \end{cases}$$

Periodic constant blur on a 3.5-ray disc + Gauss. noise with sd. 2

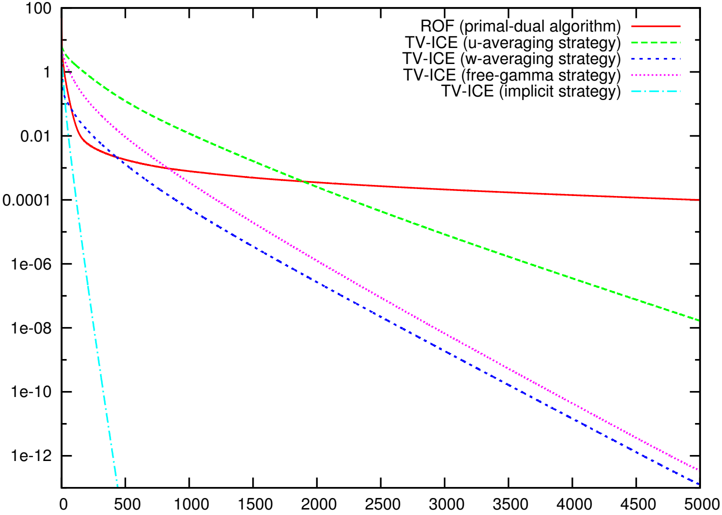
ROF deblurring

TV-ICE deblurring by averaging on u ($\gamma \approx 209$)

TV-ICE deblurring by averaging on w ($\gamma \approx 209$)

TV-ICE deblurring by free-gamma ($\gamma \approx 209$)

TV-ICE deblurring by implicit scheme ($\gamma \approx 209$)



Convergence curves

Application to zooming (from block-averaging)

Framework:

Let z be a zoom factor ($z \in \mathbb{N}^*$).
A = averaging on $z \times z$-blocks + subsampling by factor z:

$$Au(i,j) = \frac{1}{z^2}\sum_{k=0}^{z-1}\sum_{l=0}^{z-1} u(zi+k, zj+l).$$

A^*A = averaging on $z \times z$-blocks with no subsampling.

Remark: As $(I+\gamma A^*A)^{-1}(u+\gamma A^*v) = (I - \frac{\gamma}{1+\gamma}A^*A)u + \frac{\gamma}{1+\gamma}A^*v$,
3rd and 4th strategies are equivalent when $\tau = \frac{\gamma}{1+\gamma} < 2$.

Algorithm:

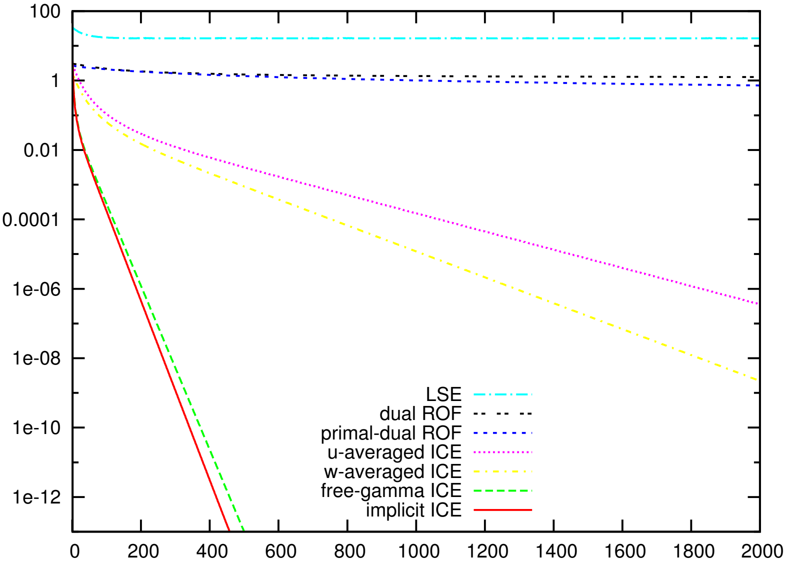
$$\begin{cases} \forall x,\ \ w^{n+1}(x) = u^n(x) + \frac{\gamma}{1+\gamma}(v(x/z) - \bar{u}^n_x) \\ \forall x,\ \ u^{n+1}(x) = F_{w^{n+1}}(u^n). \end{cases}$$

zero-order interpolation (4×4 zooming)

bicubic interpolation (4×4 zooming)

ROF (4×4 zooming)

TV-ICE (4×4 zooming)



$\lambda = 1$ and $\sigma = 5$ for each algorithm (2×2 zooming).

Outline

1. TV denoising with Iterated Conditional Expectations

2. Other (imaging?) tasks with ICE
   - Deblurring and inverse problems regularized with TV
   - TV-ICE denoising for Poisson noise
   - ICE of a convex functional
   - ICE of a convex set

TV Poisson denoising

[R. Abergel, C.L., L. Moisan, T. Zeng, SSVM 2015]

Poisson noise modelling

$$P(v|u) = \prod_{x\in\Omega} \frac{u(x)^{v(x)}}{v(x)!} \propto \exp(-\langle u - v\log u, \mathbb{1}_\Omega\rangle)$$

So TV denoising posterior p.d.f. is written as

$$P(u|v) = e^{-\langle u - v\log u, \mathbb{1}_\Omega\rangle - \lambda TV(u)}.$$

$$\begin{cases} \hat{u}_{\mathrm{MAP}} = \arg\min_u \langle u - v\log u, \mathbb{1}_\Omega\rangle + \lambda TV(u) \\ \hat{u}_{\mathrm{LSE}}(x) = \dfrac{\int_{(\mathbb{R}+)^\Omega} u(x) e^{-\langle u - v\log u, \mathbb{1}_\Omega\rangle - \lambda TV(u)}\, du}{\int_{(\mathbb{R}+)^\Omega} u(x) e^{-\langle u - v\log u, \mathbb{1}_\Omega\rangle - \lambda TV(u)}\, du} \\ \hat{u}_{\mathrm{ICE}} = \lim u^n \text{ where } u^{n+1}(x) = \dfrac{\int_{\mathbb{R}+} s^{v(x)+1} e^{-(s+\lambda\sum_{y\in\mathcal{N}_x}|s-u^n(y)|)}\, ds}{\int_{\mathbb{R}+} s^{v(x)} e^{-(s+\lambda\sum_{y\in\mathcal{N}_x}|s-u^n(y)|)}\, ds} \end{cases}$$

One ICE iteration: closed formula

Closed form

$$u^{n+1}(x) = \frac{\sum_{1\leq k\leq 5} c_k I_{a_{k-1},a_k}^{\mu_k, v(x)+1}}{\sum_{1\leq k\leq 5} c_k I_{a_{k-1},a_k}^{\mu_k, v(x)}} ,$$

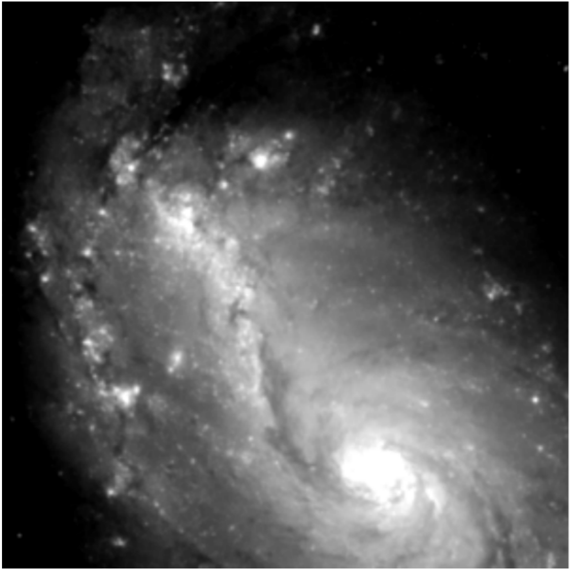
where $0 = a_0 \leq a_1 \leq a_2 \leq a_3 \leq a_3 \leq a_5 = +\infty$ with $\{a_1, a_2, a_3, a_4\} = \{u^n(\mathcal{N}_x)\}$,

$$\mu_k = 1 - (6-2k)\lambda, \quad c_k = e^{\lambda\left(\sum_{j=1}^{k-1} a_j - \sum_{j=k}^{4} a_j\right)} ,$$

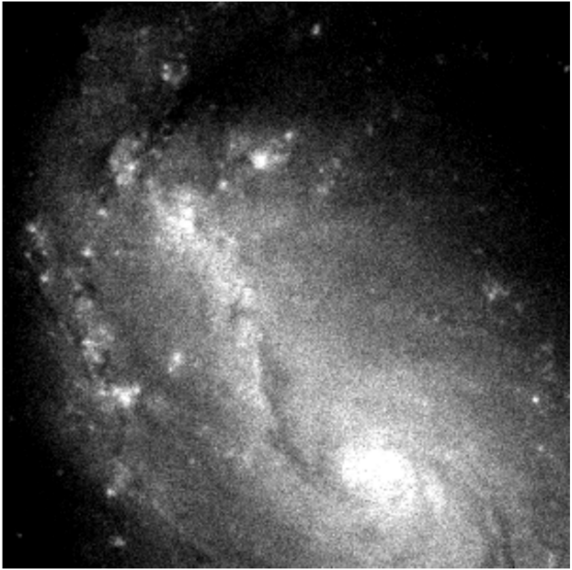
$$\text{and } I_{x,y}^{\mu,p} = \int_x^y s^p e^{-\mu s} ds .$$

It "suffices" to have good computation schemes of upper generalized gamma function and incomplete Gamma function

$$\gamma_\mu(p,x) = \int_0^x s^p e^{-\mu s} ds, \quad \Gamma_\mu(p,x) = \int_x^{+\infty} s^p e^{-\mu s} ds.$$

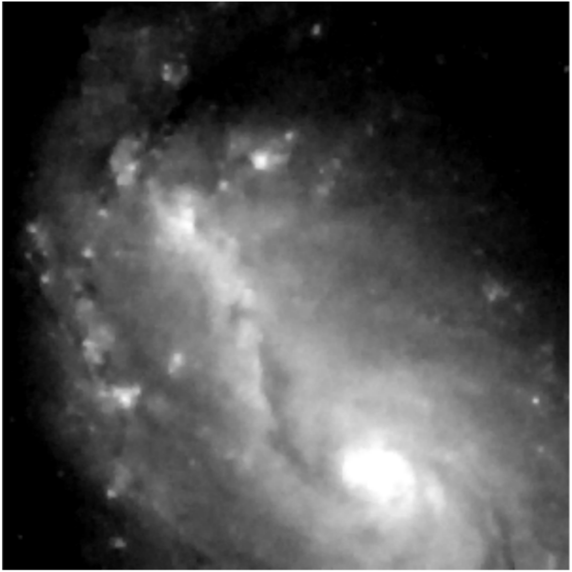


reference



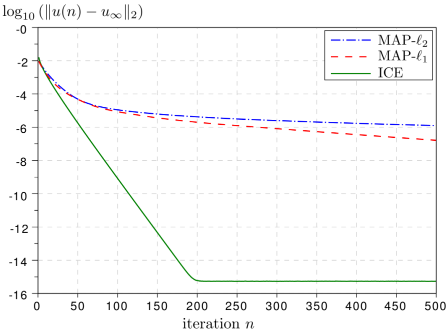
noisy

MAP



ICE

Convergence + no-staircasing



Theorem

The sequence (u^n) started at $u^0 = 0$ converges linearly to $\hat{u}_{\mathrm{ICE}}$.

No-staircasing result

Let x and $y \in \Omega$. If $\hat{u}_{\mathrm{ICE}}$ is constant on $\mathcal{N}_x \cup \mathcal{N}_y \cup \{x, y\}$, then $v(x) = v(y)$.

Outline

1. TV denoising with Iterated Conditional Expectations
2. Other (imaging?) tasks with ICE
   - Deblurring and inverse problems regularized with TV
   - TV-ICE denoising for Poisson noise
   - ICE of a convex functional
   - ICE of a convex set

Definition of ICE for a convex functional J

Let $J : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be a convex, l.s.c. and coercive function with nonempty domain $(\mathrm{dom}(J) := \{J < +\infty\})$.

Definition

An ICE point of J is a point $x \in \mathrm{dom}(J)$ such that

$$\forall 1 \leq i \leq d, \;\; x_i = \frac{\int_{\mathbb{R}} x_i e^{-J(x)} dx_i}{\int_{\mathbb{R}} e^{-J(x)} dx_i}.$$

The ICE algorithm is a sequence $(x^n)_{n\in\mathbb{N}}$ started at $x^0 \in \mathrm{dom}(J)$ and such that

$$\forall 1 \leq i \leq d, \forall n \in \mathbb{N}, \;\; x_i^{n+1} = \frac{\int_{\mathbb{R}} x_i^n e^{-J(x^n)} dx_i^n}{\int_{\mathbb{R}} e^{-J(x^n)} dx_i^n}$$

Uniqueness? Existence? Convergence?

Sufficient conditions for convergence

Theorem

1. If J is $\mathcal{C}^2$ with $\mathrm{dom}(J) = \mathbb{R}^d$, and if its Hessian H is uniformly strictly diagonally dominant, then the ICE point exists, is unique and the ICE algorithm converges linearly.
2. If J depends on an image v (so $J_v \leftarrow J$), and if
   - $u \mapsto J_v(u)$ and $v \mapsto J_v(u)$ are subdifferentiable;
   - for every $1 \leq i, j \leq d$, $\check{u}_i \in \mathbb{R}^{d-1}$, $v \in \mathbb{R}^d$, we have $u_i \mapsto \partial J_v / \partial u_j$ and $u_i \mapsto \partial J_v / \partial v_j$ nonincreasing;
   - for all $u \in \mathbb{R}^d$ and $c \in \mathbb{R}$, $J_{v+c}(u+c) = J_v(u)$;
   - for all u and $v \in \mathbb{R}^d$, $J_{-v}(-u) = J_v(u)$;
   - for all $u \geq 0$ and $v \geq 0$ in $\mathbb{R}^d$, $J_v(-u) \geq J_v(u)$;

   then the ICE point exists, is unique and the ICE algorithm converges linearly.

Example 1: $J(x) = \frac{1}{2} x^T H x - bx \Rightarrow$ ICE point=min J and ICE algo $=$ gradient descent with Jacobi preconditioning $\Rightarrow$ converges when H is strictly diag. dominant.

Sufficient conditions for convergence

Theorem

1. If J is $\mathcal{C}^2$ with $\mathrm{dom}(J) = \mathbb{R}^d$, and if its Hessian H is uniformly strictly diagonally dominant, then the ICE point exists, is unique and the ICE algorithm converges linearly.
2. If J depends on an image v (so $J_v \leftarrow J$), and if
   - $u \mapsto J_v(u)$ and $v \mapsto J_v(u)$ are subdifferentiable;
   - for every $1 \leq i, j \leq d$, $\check{u}_i \in \mathbb{R}^{d-1}$, $v \in \mathbb{R}^d$, we have $u_i \mapsto \partial J_v / \partial u_j$ and $u_i \mapsto \partial J_v / \partial v_j$ nonincreasing;
   - for all $u \in \mathbb{R}^d$ and $c \in \mathbb{R}$, $J_{v+c}(u + c) = J_v(u)$;
   - for all u and $v \in \mathbb{R}^d$, $J_{-v}(-u) = J_v(u)$;
   - for all $u \geq 0$ and $v \geq 0$ in $\mathbb{R}^d$, $J_v(-u) \geq J_v(u)$;

   then the ICE point exists, is unique and the ICE algorithm converges linearly.

Example 2: $J_v(u) = \text{Poisson noise}(v, u) + TV(u) \Rightarrow$ linear convergence.

Outline

1. TV denoising with Iterated Conditional Expectations

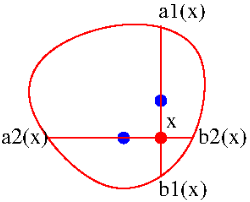
2. Other (imaging?) tasks with ICE
   - Deblurring and inverse problems regularized with TV
   - TV-ICE denoising for Poisson noise
   - ICE of a convex functional
   - ICE of a convex set

Definition of ICE for a convex set C

Consider the canonical basis $(e_i)_{1\leq i\leq d}$ of $\mathbb{R}^d$ (or another basis). Let $C \subset \mathbb{R}^d$ be a nonempty compact convex set.

Definition

An ICE point of C (relatively to the basis (e_i)) is a point $x \in C$ that is the midpoint of $[a_i(x), b_i(x)]$ for each $1 \leq i \leq d$, where $[a_i(x), b_i(x)] := C \cap \{x + te_i, t \in \mathbb{R}\}$.
The ICE algorithm is a sequence $(x^n)_{n\in\mathbb{N}}$ started at $x^0 \in C$ and such that x^{n+1} is the middle of $[a_i(x^n), b_i(x^n)]$ $(1 \leq i \leq d)$.

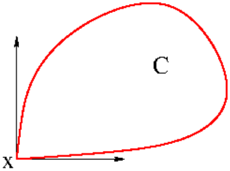


- existence of an ICE point by Schauder fixed-point theorem;
- a rectangle may have infinitely many ICE points;
- counterpart for center of gravity;
- translation- but not rotation-invariant.

Definition

An extremal point of C is a point $x \in C$ such that

$$\forall 1 \leq i \leq d,\ \forall y, z \in C \setminus \{x\},\ \langle y - x, e_i \rangle \cdot \langle z - x, e_i \rangle > 0.$$



We have x extremal point $\Rightarrow$ x ICE point.
$\Rightarrow$ even (strictly) convex sets may have several ICE points
(e.g. $B((1,0),1) \cap B((0,1),1) \subset \mathbb{R}^2$).

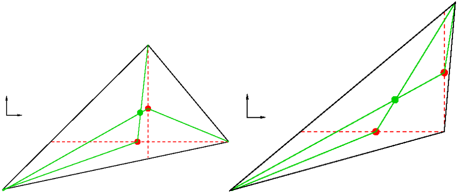
Observation

1. If C is strictly convex and has no extremal point (e.g. if C has a $\mathcal{C}^2$ boundary), then the ICE point is unique.
2. If C has a $\mathcal{C}^2$ boundary then the algorithm converges to an ICE point and the convergence is linear.

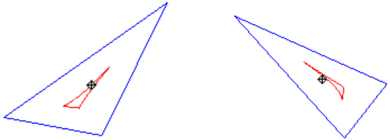
Case of the triangle

Theorem

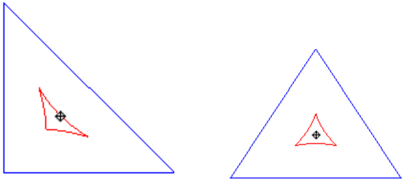
1. A triangle has (apart from its extremal vertices) a unique ICE point.
2. The ICE algorithm (**not** initialized on an extremal vertex) converges linearly.
3. ICE point found by a simple geometric construction.



Construction of ICE



Location of ICE ∀ rotation



right isosceles   equilateral

Concluding remarks

- framework for some image restoration problems without energy minimization
- fast approximation of LSE
- purely primal algorithm: no big theory + any initialization + nice-to-see convergence
- raises interesting questions even in low dimension.