Karlsruhe Institute of Technology

A case study of the use of discrete gradient methods in image processing

Volker Grimm (joint work with R. McLachlan, D. McLaren, G. R. W. Quispel, C.-B. Schönlieb)

Institute for Applied and Numerical Mathematics 1

Outline

Gradient flows

- Gradient flow in Hilbert spaces

Discrete gradient methods

- Discrete gradients

Nonlinear problems in image processing

- Total variation deblurring
- Multichannel TV denoising

Gradient flow in Hilbert spaces

Let H be a Hilbert space with inner product $\langle \cdot\,, \cdot \rangle$ and let $V\,:\,H \to \mathbb{R}$ be a differentiable functional.

The gradient of V at $x \in H$ is the unique element $\nabla V(x)$ satisfying

$$\langle \nabla V(x), v \rangle = \frac{d}{dt} V(x + tv)\Big|_{t=0} \quad \text{for all} \quad v \in H.$$

A gradient flow is the solution of the initial value problem

$$\dot{x} = -\nabla V(x), \qquad x(0) = x_0$$

with decay

$$\frac{d}{dt} V(x(t)) = -\|\nabla V(x(t))\|^2 \leq 0.$$

Gradient systems in image processing

Karlsruhe Institute of Technology

- TV regularisation
- time-marching schemes (e.g. find solution of Tikhonov regularisation)
- Perona-Malik model and many variants
- Sobolev gradient flows, general metrics
- enhancing of images (smoothing, sharpening,...)
- registration of two and more images
- super resolution
- snakes
- level sets
- nonlinear diffusion filters
- diffusion process as regularisation

Discrete gradients

Let $V : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable.

Then $\overline{\nabla} V : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ is a discrete gradient of V
iff it is continuous and

$$\begin{cases} \langle \overline{\nabla} V(x, x'), (x' - x) \rangle = V(x') - V(x), \\ \overline{\nabla} V(x, x) = \nabla V(x) \end{cases} \quad \text{for all} \quad x, x' \in \mathbb{R}^n.$$

Capel, Celledoni, Cohen, Furihata, Gonzales, Hairer, Lubich, Matsuo, McLaren, McLachlan, O'Neale, Owren, Quispel, Robidoux, Schönlieb, Stuart, Turner, Wright, ...

Discrete gradients

- midpoint discrete gradient / Gonzales discrete gradient $(x \neq x')$

$$
\begin{aligned}
\overline{\nabla}_1 V(x,x') \quad &= \quad \nabla V\left(\frac{x'+x}{2}\right) \\
&+ \frac{V(x') - V(x) - \left\langle \nabla V\left(\frac{x'+x}{2}\right), (x'-x)\right\rangle}{\|x - x'\|^2}(x'-x)
\end{aligned}
$$

- mean value discrete gradient

$$
\overline{\nabla}_2 V(x,x') = \int_0^1 \nabla V((1-s)x + sx')\,ds
$$

Discrete gradient method

For the gradient flow

$$\dot{x} = -\nabla V(x), \qquad x(0) = x_0,$$

every discrete gradient $\overline{\nabla} V$ leads to an associated discrete gradient method

$$x_{n+1} - x_n = -\tau_n \overline{\nabla} V(x_n, x_{n+1}) \,.$$

Preservation of decay

$$\begin{aligned} V(x_{n+1}) - V(x_n) &= \langle \overline{\nabla} V(x_n, x_{n+1}), (x_{n+1} - x_n) \rangle \\ &= -\tau_n \|\overline{\nabla} V(x_n, x_{n+1})\|^2 \\ &\leq 0 \end{aligned}$$

Convergence to minimizers

Theorem

Let ∇V stem from a functional V which is bounded from below, coercive and continuously differentiable. If $\{x_n\}_{n=0}^{\infty}$ is a sequence generated by the discrete gradient method with time steps $0 < c \leq \tau_n \leq M < \infty$.

1. If V is in addition convex, then a minimizer exists and any accumulation point of the sequence $\{x_n\}_{n=0}^{\infty}$ is a minimizer.
2. If V is in addition strictly convex, then

$$\lim_{n\to\infty} x_n = x_*, \qquad V(x_*) = \min_x V(x) .$$

Since V is bounded from below

$$C \leq V(x_{n+1}) \leq V(x_n) \leq \cdots \leq V(x_0), \qquad n = 1, 2, 3, \ldots$$

and hence the limit

$$\lim_{n \to \infty} V(x_n) = V_*$$

exists. From the definition of the discrete gradient we find

$$\begin{aligned}
\tau_n \|\overline{\nabla} V(x_{n+1}, x_n)\|^2 &= -\langle \overline{\nabla} V(x_{n+1}, x_n), x_{n+1} - x_n \rangle = V(x_n) - V(x_{n+1}) \\
&= \frac{1}{\tau_n} \langle -\tau_n \overline{\nabla} V(x_{n+1}, x_n), x_{n+1} - x_n \rangle \\
&= \frac{1}{\tau_n} \|x_{n+1} - x_n\|^2 \geq 0
\end{aligned}$$

By summing these equations from n to $m-1$, $m>n$, we obtain

$$\sum_{k=n}^{m-1} \tau_k \left\|\overline{\nabla} V(x_{k+1}, x_k)\right\|^2 = \sum_{k=n}^{m-1} \frac{1}{\tau_k} \|x_{k+1} - x_k\|^2 \leq V(x_0) - V_*$$

and thus

$$\begin{aligned} \sum_{k=0}^{\infty} \left\|\overline{\nabla} V(x_{k+1}, x_k)\right\|^2 &\leq \frac{V(x_0) - V_*}{c} < \infty \\ \sum_{k=0}^{\infty} \|x_{k+1} - x_k\|^2 &\leq M\left(V(x_0) - V_*\right) < \infty \end{aligned}$$

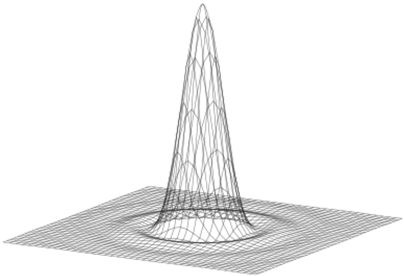
and therefore

$$\lim_{n\to\infty} (x_{n+1} - x_n) = \lim_{n\to\infty} \overline{\nabla} V(x_{n+1}, x_n) = 0\,.$$

Total Variation (TV) deblurring

Functional

$$T_\alpha(u) = \frac{1}{2}\int_\Omega \big((Ku)(x,y) - u_0(x,y)\big)^2\, d(x,y) + \alpha \mathrm{TV}(u)$$



Total Variation (TV) deblurring

Functional

$$T_\alpha(u) = \frac{1}{2}\int_\Omega \big((Ku)(x,y) - u_0(x,y)\big)^2\, d(x,y) + \alpha \mathrm{TV}(u)$$

TV functional

$$\mathrm{TV}(u) = \int_\Omega \sqrt{\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \beta}\; d(x,y)$$

Parabolic gradient system

$$u_t = -\nabla T_\alpha(u) = \alpha \nabla \cdot \left[\frac{\nabla u}{|\nabla u|_\beta}\right] - K^*(Ku - u_0) \qquad \text{with} \qquad \left.\frac{\partial u}{\partial \mathbf{n}}\right|_{\partial\Omega} = 0$$

Discretised Energy

Discretised functional

$$V_\alpha(u) = \frac{1}{2}\Delta x \Delta y \left\| \tilde{K} u - u_0 \right\|^2 + \alpha J(u)$$

Discretised TV functional

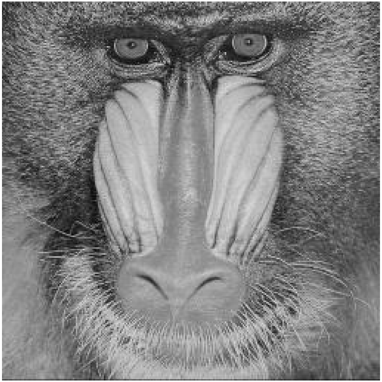
$$J(u) = \Delta x \Delta y \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} \psi\left((D_{ij}^x u)^2 + (D_{ij}^y u)^2 \right),$$

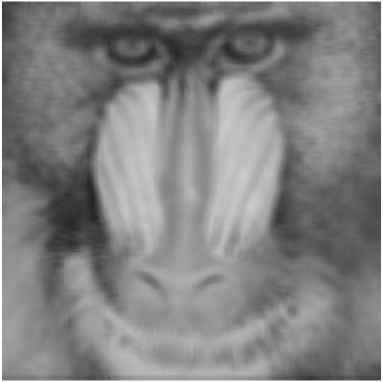
where $\psi(t) = \sqrt{t + \beta}$ and

$$D_{ij}^x u = \frac{u_{i,j} - u_{i-1,j}}{\Delta x}, \qquad D_{ij}^y u = \frac{u_{i,j} - u_{i,j-1}}{\Delta y}$$

TV deblurring

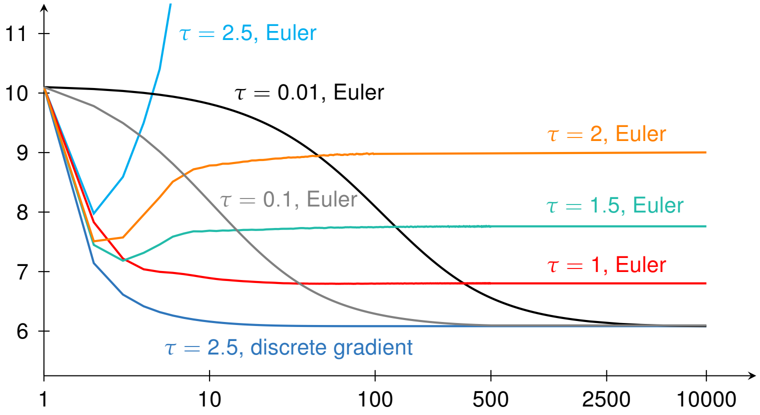
Original image and blurred image





Functional decay

Karlsruhe Institute of Technology



TV deblurring with discrete gradient method

Large step sizes

- Newton method in inner iteration
- exact Jacobian
- preconditioned CG method

Small step sizes

- fixed-point iteration
- explicit method

Multichannel TV denoising

Multichannel TV functional

$$\mathrm{TV}_2[u] = \left(\sum_{i=1}^{p} \left(\mathrm{TV}[u_i]\right)^2 \right)^{1/2} = \left(\sum_{i=1}^{p} \left(\int_\Omega |Du_i| \, d(x,y) \right)^2 \right)^{1/2}$$

Multichannel functional, p channels

$$T_\alpha(u) = \alpha \mathrm{TV}_2[u] + \frac{1}{2} \int_\Omega \|u - u_0\|^2 \, d(x,y) \, .$$

With the global constants

$$c_i[u] = \frac{\mathrm{TV}[u_i]}{\mathrm{TV}_2[u]} \geq 0, \qquad i = 1, \ldots, p \, ,$$

we have the gradient system

$$\frac{d}{dt} u_i = \alpha \cdot c_i[u] \nabla \cdot \left[\frac{\nabla u_i}{|\nabla u_i|_\beta} \right] - (u_i - u_{0,i}) = 0, \quad \left. \frac{\partial u_i}{\partial \mathbf{n}} \right|_{\partial\Omega} = 0, \quad i = 1, \ldots, p \, .$$

Macro photography

Picture data

- Canon MP-E 65mm macro lens
- extremely low depth-of-field
- Canon EOS 550D
- hand-held in full sunlight
- exposure time $1/250$
- f-stop number 14
- 3x magnification
- film speed ISO 6400

Problem: High film speed produces a lot of noise due to amplification of the signal from the charge-coupled device (CCD) image sensor.

TV denoising

Original image

TV denoising

Denoised image

TV denoising

Detail of original image

TV denoising

Detail of denoised image

Conclusion

Discrete gradient (DG) methods

- DG methods form a broad class of methods
- DG methods preserve gradient structure

DG methods in image processing

- DG method allows for larger step sizes
- DG methods applicable to nonlinear problems
- DG works where other implicit methods do not work
- DG more reliable, even for small step sizes
- DG in real world problems, only α needs to be tuned
- DG methods better for automatic application