Bregman iteration for optical flow

Laurent Hoeltgen

18th December 2009

Outline

Constrained optimisation problems

- Let J and H be two convex functions from $\mathbb{R}^n$ to $\mathbb{R}$
- Assume $\min_u H(u) = 0$
- Consider the constrained problem

$$\arg\min_u \; J(u) \quad \text{such that} \quad H(u) = 0$$

- Can be very difficult to solve.
  - $H(u) = 0$ can have infinitely many solutions.
  - J and H not necessarily differentiable.

Usual approach

- Constrained optimization problem

$$\arg\min_u \; J(u) \quad \text{such that} \quad H(u) = 0$$

- Approximate solution by solving series of problems

$$\arg\min_u \; J(u) + \lambda_n H(u)$$

for $0 < \lambda_1 < \lambda_2 < \lambda_3 < \ldots < \lambda_N$.

- Approach is called penalty function/continuation method.

Disadvantages

- Requires very large λ_N for good approximation.
  - Numerically unstable/ill-conditioned for large λ_N.
- Sometimes the λ_i can only be increased in small steps.
  - Algorithm becomes slow due to lots of computations.
- Can be as difficult to solve as initial problem.

Are there alternative approaches without these problems?

- Bregman iteration presents interesting alternative.

Subdifferential and subgradient

Assume $f : \mathbb{R}^n \to \mathbb{R}$ is a convex function.

Definition

We call subdifferential of f at y the set

$$\partial f(y) := \{q \in \mathbb{R}^n : f(x) \geqslant f(y) + \langle q, x - y \rangle, \ \forall x \in \mathbb{R}^n\}$$

Its elements are called subgradients.

Example

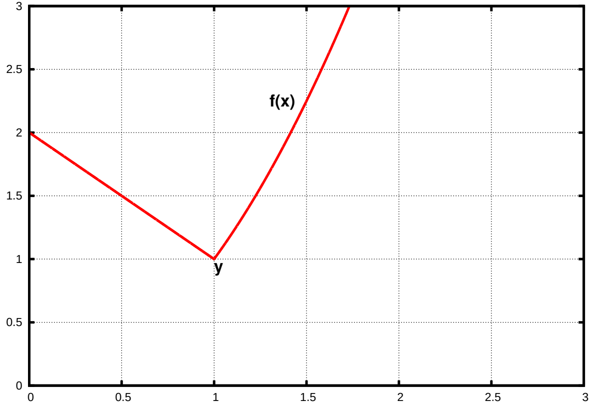


Figure: A convex function f

Example

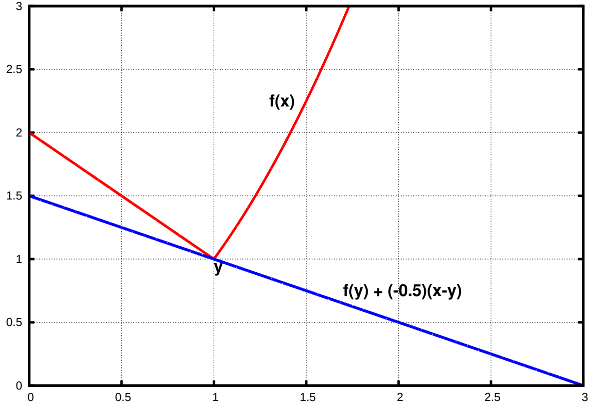


Figure: q$=-\frac{1}{2}$ is a subgradient of f at y

Example

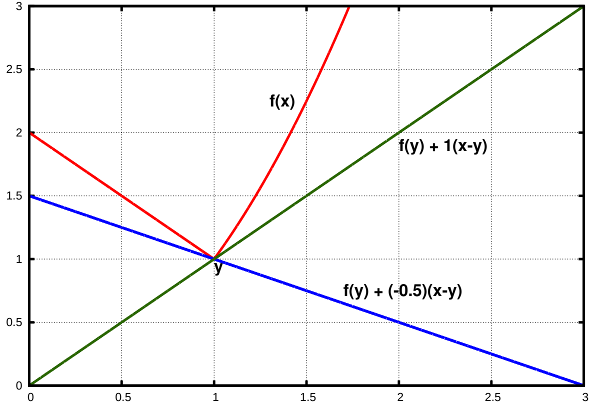


Figure: q=1 another subgradient of f at y

Bregman divergence

Again assume f is a convex function.

Definition

We call Bregman divergence of f the function:

$$D_f^q(x, y) := f(x) - f(y) - \langle q, x - y \rangle$$

- y is an arbitrary but fixed point.
- q is a subgradient of f at y.

Interpretation

- Consider convex and differentiable function $f : \mathbb{R}^n \to \mathbb{R}$
- Linearising f with Taylor expansion around y gives

$$L_{f,y}(x) := f(y) + \langle \nabla f(y), x - y \rangle$$

- Difference between $f(x)$ and $L_{f,y}(x)$ is

$$f(x) - L_{f,y}(x) = f(x) - f(y) - \langle \nabla f(y), x - y \rangle$$

- This is the definition of the Bregman divergence!

Definition of the Bregman iteration

- Bregman iteration is an algorithm for solving

$$\arg\min_u \; J(u) \quad \text{such that} \quad H(u) = 0$$

for convex J and H from $\mathbb{R}^n$ to $\mathbb{R}$ and $\min_u H(u) = 0$

- It computes iteratively

$$\begin{aligned} u^{k+1} &= \arg\min_u \; D_J^{p^k}\left(u, u^k\right) + \lambda H(u) \\ &= \arg\min_u \; J(u) - J\left(u^k\right) - \langle p^k, u - u^k \rangle + \lambda H(u) \\ &= \arg\min_u \; J(u) - \langle p^k, u - u^k \rangle + \lambda H(u) \end{aligned}$$

with $p^k \in \partial J\left(u^k\right)$ and arbitrary $\lambda > 0$.

Advantages of Bregman iteration

At first glance similar to penalty function methods.

However

- $\lambda > 0$ is arbitrary and does not need to be large.
  - Numerically much more stable.
- Only 1 problem to solve instead of series of problems.
  - Faster than penalty function/continuation methods.
- Convergence towards solution of constrained problem.

Convergence results

If $H(u) := \|Au - b\|_2^2$ with matrix A and vector b.

Then one can show

- $0 \leqslant H\left(u^{k+1}\right) \leqslant H\left(u^k\right)$
- If $H\left(u^k\right) = 0$, then u^k also solves

$$\arg\min_u \; J(u) \quad \text{such that} \quad H(u) = 0$$

Preliminary conclusions

- Constrained optimisation problems are difficult to solve.
- Conventional approaches may have disadvantages.
- Bregman iteration is an interesting alternative.

Applying Bregman iteration in optical flow

Simple model:

- Grey value constancy

$$f\,(x+u, y+v, t) = f\,(x, y, t+1)$$

- Linearised constancy assumption.

$$f_x \cdot u + f_y \cdot v + f_t = 0$$

- Flow field should also be smooth (smoothness constraint)

$$\|\nabla u\|_1 + \|\nabla v\|_1 \text{ should be small}$$

Problem formulation

- Find minimizer (u, v) of

$$\arg\min_{u,v} \|\nabla u\|_1 + \|\nabla v\|_1 + \frac{\mu}{2}\|f_x \cdot u + f_y \cdot v + f_t\|_2^2$$

- Difficult to solve: $\|\cdot\|_1$ is not differentiable.

Problem formulation

- Find minimizer (u, v) of

$$\arg\min_{u,v} \|\nabla u\|_1 + \|\nabla v\|_1 + \frac{\mu}{2}\|f_x \cdot u + f_y \cdot v + f_t\|_2^2$$

- Define

$$w := (u, v)^T$$
$$F\, w := f_x \cdot u + f_y \cdot v$$

- Problem rewrites as

$$\arg\min_{w} \|\nabla w\|_1 + \frac{\mu}{2}\|Fw + f_t\|_2^2$$

Problem formulation

Unconstrained problem

$$\arg\min_{w} \|\nabla w\|_1 + \frac{\mu}{2}\|Fw + f_t\|_2^2$$

can be rewritten as constrained problem

$$\arg\min_{w,d} \|d\|_1 + \frac{\mu}{2}\|Fw + f_t\|_2^2 \quad \text{such that} \quad d = \nabla w$$

Problem formulation

- Find minimizer (w, d) of

$$\arg\min_{w,d} \|d\|_1 + \frac{\mu}{2}\|Fw + f_t\|_2^2 \quad \text{such that} \quad d = \nabla w$$

- Define

$$\begin{aligned} \eta &:= (w, d) \\ J(\eta) &:= \|d\|_1 + \frac{\mu}{2}\|Fw + f_t\|_2^2 \\ A(\eta) &:= d - \nabla w \end{aligned}$$

- Problem rewrites as

$$\arg\min_{\eta} \; J(\eta) \quad \text{such that} \quad \|A\eta\|_2^2 = 0$$

Solving through Bregman iteration

- Find minimizer η of

$$\arg\min_{\eta} \; J(\eta) \quad \text{such that} \quad \|A\,\eta\|_2^2 = 0$$

- Apply Bregman iteration:

$$\eta^{k+1} = \arg\min_{\eta} \; D_J^{p^k}\left(\eta, \eta^k\right) + \frac{\lambda}{2}\|A\,\eta\|_2^2$$
$$p^k \in \partial J\left(\eta^k\right)$$

- One can show following simplification is possible:

$$\eta^{k+1} = \arg\min_{\eta} \; J(\eta) + \frac{\lambda}{2}\|A\,\eta - b^k\|_2^2$$
$$b^{k+1} = b^k - A\eta^{k+1}$$

Solving through Bregman iteration

- Leads to this algorithm

$$\underbrace{\left(w^{k+1}, d^{k+1}\right)}_{=\,\eta^{k+1}} = \arg\min_{w,d} \underbrace{\|d\|_1 + \frac{\mu}{2}\|Fw + f_t\|_2^2}_{=\,J(\eta)} + \frac{\lambda}{2}\|\underbrace{d - \nabla w}_{=\,A\eta} - b^k\|_2^2$$

$$b^{k+1} = b^k - \left(d^{k+1} - \nabla w^{k+1}\right)$$

- Easily solvable through simple 3 step algorithm:

Step 1: $w^{k+1} = \arg\min_{w} \frac{\mu}{2}\|Fw + f_t\|_2^2 + \frac{\lambda}{2}\|d^k - \nabla w - b^k\|_2^2$

Step 2: $d^{k+1} = \arg\min_{d} \|d\|_1 + \frac{\lambda}{2}\|d - \nabla w^{k+1} - b^k\|_2^2$

Step 3: $b^{k+1} = b^k - \left(d^{k+1} - \nabla w^{k+1}\right)$

Solving step 1

$$w^{k+1} = \arg\min_{w} \ \frac{\mu}{2}\|Fw + f_t\|_2^2 + \frac{\lambda}{2}\|d^k - \nabla w - b^k\|_2^2$$

- Cost function is differentiable. Requires solving linear system.
  - Matrix is positive definite, sparse and very large.
  - Use Gauss-Seidel or similar algorithms.

Solving step 2

$$d^{k+1} = \arg\min_{d} \|d\|_1 + \frac{\lambda}{2}\|d - \nabla w^{k+1} - b^k\|_2^2$$

- Step 2 has analytical solution:

$$d_j^{k+1} = \text{shrink}\left(\left(\nabla w^{k+1} + b^k\right)_j, \frac{1}{\lambda}\right) \quad j = 1, \dots, n$$

where $\text{shrink}(x, \gamma) := \frac{x}{|x|} \max(|x| - \gamma, 0)$.

  - Operates componentwise.
  - Only simple operations. Very fast.

Solving step 3

$$b^{k+1} = b^k - \left(d^{k+1} - \nabla w^{k+1}\right)$$

- Step 3 is just a simple computation.

Example

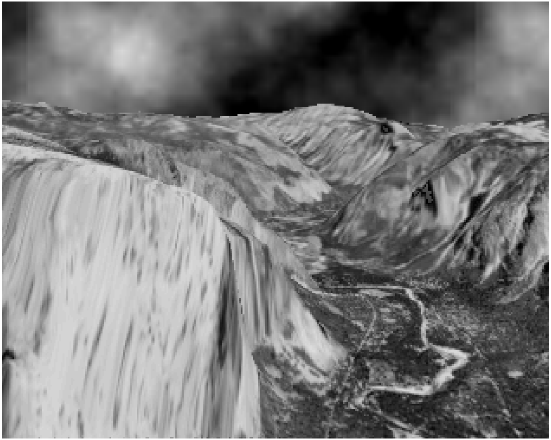


Figure: Yosemite sequence frame 10

Example

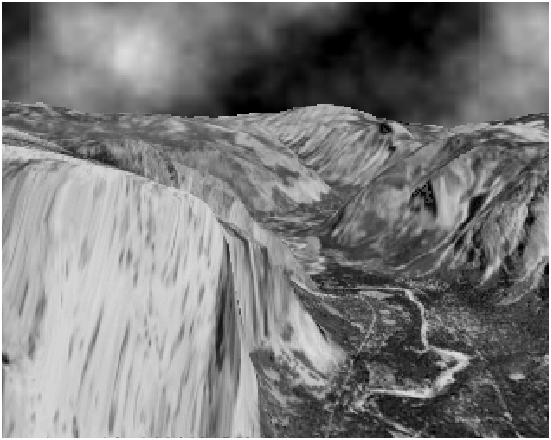
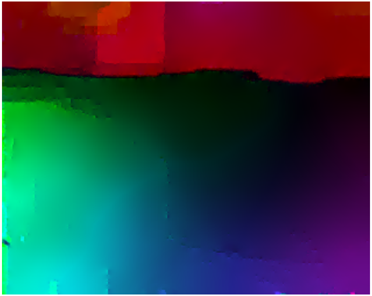
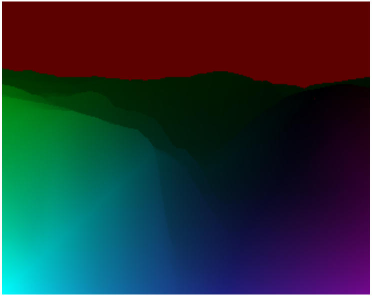


Figure: Yosemite sequence frame 11

Example



(a) Result



(b) Exact solution

Figure: Qualitative comparison: Bregman result and exact ground truth

Conclusion

+ Results look as expected.
  - Average angular error of 5.65.
+ Only 65 Iterations, 5 Iterations for Gauss-Seidel each time.
  - Algorithm is very fast.
+ Very simple implementation.

– 4 different parameters. Not clear how to chose them.
  - In this case: $\lambda = 36.52$, $\mu = 0.49$, 65 and 5 iterations

+ Useful for functionals that are difficult to minimize.

Aims of the thesis

- Find good models for the Bregman Iteration.
- Theoretical and numerical evaluation
  - Convergence behaviour
  - Comparison to other modern approaches

Bibliography

- S. Osher, M. Burger, D. Goldfarb, J. Xu and W. Yin
  An iterative regularization method for total variation image restoration.
  UCLA CAM Report 04-13.

- S. Osher, M. Burger, D. Goldfarb and J. Darbon
  Bregman iterative algorithm for ℓ_1 minimization with applications to compressed sensing.
  UCLA CAM Report 07-37.

- S. Osher and T. Goldstein
  The split Bregman Method for ℓ_1 regularized problems.
  UCLA CAM Report 08-29.

Bibliography

L.M. Bregman
The relaxation method for finding the common point of convex sets and its application to the solution of problems in convex programming.
USSR Computational Mathematics and Mathematical Physics, 7:200-217,1967

M. Burger, E. Resmerita and L. He
Error estimation for Bregman iterations and inverse scale space methods.
Computing, 81:109-135, 2007

Thank you for your attention.