Variational Methods in Biomedical Imaging
Part II: Optimisation

Luis Pizarro

Biomedical Image Analysis Group
Department of Computing
Imperial College, London, UK

Outline

- Discretisation Strategies
- Numerical Methods
- Principles of Duality
- Convex Problems
- Summary

Variational Restoration in 1-D

The minimiser $u : [a,b] \to \mathbb{R}$ of the functional

$$E(u) := \int_a^b (u-f)^2 \, dx \; + \; \alpha \int_a^b \Psi\left(u_x^2\right) \, dx$$

necessarily satisfies the Euler-Lagrange equation

$$0 = u - f - \alpha \, \partial_x \left(\Psi'(u_x^2) \, u_x\right)$$

with the Neumann boundary conditions

$$u_x = 0 \qquad \text{for } x = a \text{ and } x = b.$$

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

Discretisation of the Euler-Lagrange Equation

- Consider the grid size $h := \frac{b-a}{N}$ and grid points $x_i := (i - \frac{1}{2})$ with $i = 1, \ldots, N$. Let u_i denote an approximation to $u(x_i)$.
- The discretised Euler-Lagrange equation for all $i = 2, \ldots, N-1$ reads

$$\begin{aligned} 0 \quad = \quad u_i - f_i \quad & - \quad \alpha \, \Psi'\left(\frac{(u_{i+1}-u_i)^2}{h^2}\right) \frac{(u_{i+1}-u_i)}{h^2} \\ & + \quad \alpha \, \Psi'\left(\frac{(u_i-u_{i-1})^2}{h^2}\right) \frac{(u_i-u_{i-1})}{h^2} \, . \end{aligned}$$

- With $\mathbf{u} = (u_1, \ldots, u_N)^\top$ and the sparse $N \times N$ matrix of coefficients $A(\mathbf{u}) = (a_{k,l}(\mathbf{u}))$, the entire system reads

$$0 = \mathbf{u} - \mathbf{f} - \alpha A(\mathbf{u})\mathbf{u} \, .$$

- Thus, we have to solve the nonlinear system of equations

$$\Big(I - \alpha A(\mathbf{u})\Big)\mathbf{u} = \mathbf{f} \, .$$

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

Discretisation of the Energy Functional

- A discrete realisation of the original continuos model reads

$$E(\mathbf{u}) := \sum_{i=1}^{N} (u_i - f_i)^2 \;+\; \alpha \sum_{i=1}^{N-1} \Psi\left(\frac{(u_{i+1}-u_i)^2}{h^2}\right) .$$

- Setting $\frac{\partial E}{\partial u_i} = 0$ for all $i = 2, \ldots, N-1$ gives

$$\begin{aligned} 0 \;=\; u_i - f_i \;&-\; \alpha\, \Psi'\left(\frac{(u_{i+1}-u_i)^2}{h^2}\right) \frac{(u_{i+1}-u_i)}{h^2} \\ &+\; \alpha\, \Psi'\left(\frac{(u_i-u_{i-1})^2}{h^2}\right) \frac{(u_i-u_{i-1})}{h^2} . \end{aligned}$$

- Same nonlinear system as before without using the Euler-Lagrange equation.
- This strategy can be beneficial if you have difficulties discretising the boundary conditions or if you want to prove stability results.

Numerical Methods for the Elliptic Problem

- The elliptic problem (Euler-Lagrange equation)

$$0 \;=\; \nabla_{\mathbf{u}} E \;=\; \mathbf{u} - \mathbf{f} - \alpha A(\mathbf{u})\mathbf{u}$$

  amounts to solving the nonlinear system

$$\Big(I - \alpha A(\mathbf{u})\Big)\mathbf{u} = \mathbf{f} .$$

- The so-called *Kačanov method* solves the nonlinear system as a sequence of linear problems (explicit scheme):

$$\Big(I - \alpha A(\mathbf{u}^{\mathbf{k}})\Big)\mathbf{u}^{k+1} = \mathbf{f} \qquad (k = 1, 2, \ldots) .$$

- It can be regarded as a fixed point iteration of:

$$\mathbf{u} = \Big(I - \alpha A(\mathbf{u})\Big)^{-1} \mathbf{f} .$$

- Use classical iterative solvers for large linear systems $B\mathbf{x} = c$ such as: Jacobi, Gauß-Seidel, Successive Overrelaxation (SOR)

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

Numerical Methods for the Parabolic Problem

- The parabolic problem comes from considering the gradient descent of the elliptic problem

$$\begin{aligned}
\frac{\partial u}{\partial t} &= -\nabla_u E \\
&= \operatorname{div}\left(\Psi'(|\nabla u|^2)\nabla u\right) - \frac{u-f}{\alpha} \\
&= \sum_{l=1}^{d} \partial_{x_l}\left(\Psi'(|\nabla u|^2)\, u_{x_l}\right) - \frac{u-f}{\alpha}
\end{aligned}$$

- We explore three numerical algorithms for this problem
  - Modified explicit-scheme
  - Semi-implicit scheme
  - Additive operator splitting (AOS) scheme

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

Numerical Methods for the Parabolic Problem

- **Modified explicit scheme**

  Explicit approximation (implicit in bias term):

$$\frac{\mathbf{u}^{k+1}-\mathbf{u}^{k}}{\tau} = \sum_{l=1}^{d} A_l^k\, \mathbf{u}^k + \frac{1}{\alpha}(\mathbf{f} - \mathbf{u}^{k+1})\ .$$

  It can be solved directly (explicitly) for the unknown $\mathbf{u}^{k+1}$

$$\mathbf{u}^{k+1} = \frac{\alpha}{\alpha+\tau}\left(I + \tau\sum_{l=1}^{d} A_l^k\right)\mathbf{u}^k + \frac{\tau}{\alpha+\tau}\mathbf{f}\ .$$

  This scheme is stable for $\tau \leq \left(\frac{1}{2}\right)^d$, which makes it relatively slow to converge.

Numerical Methods for the Parabolic Problem

◆ Semi-implicit scheme

Semi-implicit approximation:

$$\frac{\mathbf{u}^{k+1} - \mathbf{u}^k}{\tau} = \sum_{l=1}^{d} A_l^k \, \mathbf{u}^{k+1} + \frac{1}{\alpha}(\mathbf{f} - \mathbf{u}^{k+1}) \; .$$

It is absolutely stable for $\tau > 0$, but requires to solve the linear system

$$\left(I - \frac{\alpha\tau}{\alpha+\tau} \sum_{l=1}^{d} A_l^k \right) \mathbf{u}^{k+1} = \frac{\alpha\mathbf{u}^k + \tau\mathbf{f}}{\alpha+\tau} \; .$$

Appropriate solvers include Jacobi, Gauß-Seidel, SOR.

Numerical Methods

Numerical Methods for the Parabolic Problem

◆ Additive operator splitting (AOS) scheme

Replace semi-implicit scheme:

$$\mathbf{u}^{k+1} = \left(I - \frac{\alpha\tau}{\alpha+\tau} \sum_{l=1}^{d} A_l^k \right)^{-1} \frac{\alpha\mathbf{u}^k + \tau\mathbf{f}}{\alpha+\tau} \; .$$

by the absolutely stable additive operator splitting

$$\mathbf{u}^{k+1} = \frac{1}{d} \sum_{l=1}^{d} \left(I - d\,\frac{\alpha\tau}{\alpha+\tau} A_l^k \right)^{-1} \frac{\alpha\mathbf{u}^k + \tau\mathbf{f}}{\alpha+\tau} \; .$$

This splits the solution of a d-dimensional problem into the solution of d 1-dimensional problems, each of them being a tridiagonal system that can be easily solved by the Thomas algorithm.

Duality between Points and Lines

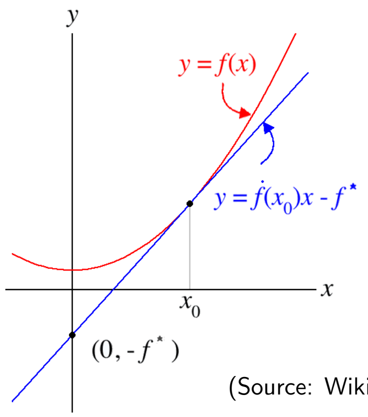
A curve can be represented as a locus of points or as an envelope of tangents. Both representations are dual to each other.

Legendre-Fenchel (LF) transformation

The convex conjugate $f^*(k)$ of a function $f(x)$ reads

$$f^*(k) = \sup_{x \in \Omega_x} \Big(k\,x - f(x) \Big) \;,$$

with the dual variable (slope) $k \in \Omega_k$.



(Source: Wikipedia)

Geometrically, find x such that the line with slope k passing through $(x, f(x))$ has a maximum intercept with the y-axis.

Principles of Duality

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

- In higher dimensions, $f : \mathbb{R}^n \to \mathbb{R}$

$$f^*(k) = \sup_{x \in \Omega_x} \Big(< k, x > - f(x) \Big)$$

- The function $f^*(k)$ is always convex, irrespective of the shape of $f(x)$

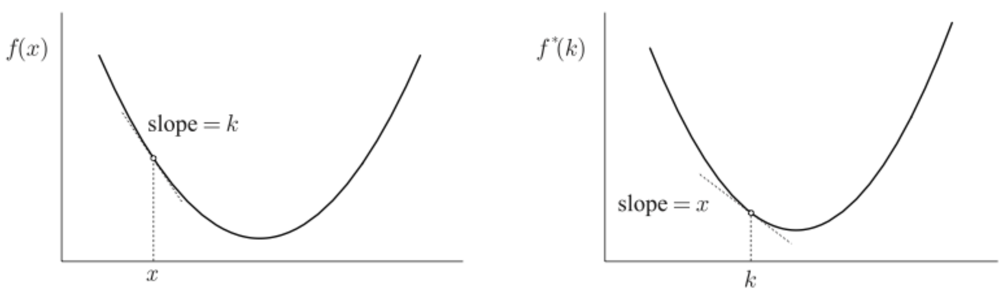
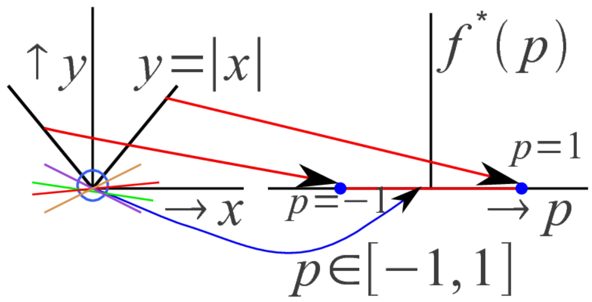


Illustration of duality: points of f are transformed into slopes of f^*, and slopes of f are transformed into points of f^*. (Source: Touchette 2007)

Principles of Duality

- The function $f(x) = |x|$, non-differentiable at $x_0 = 0$,



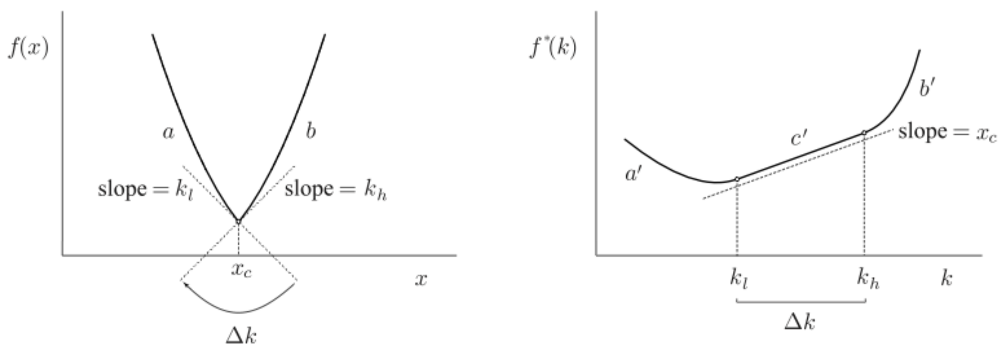
has infinite lines with slope $k \in [-1, 1]$ passing through $(x_0, f(x_0))$.

- The subdifferential of $f(x)$ at a point y is defined as the set

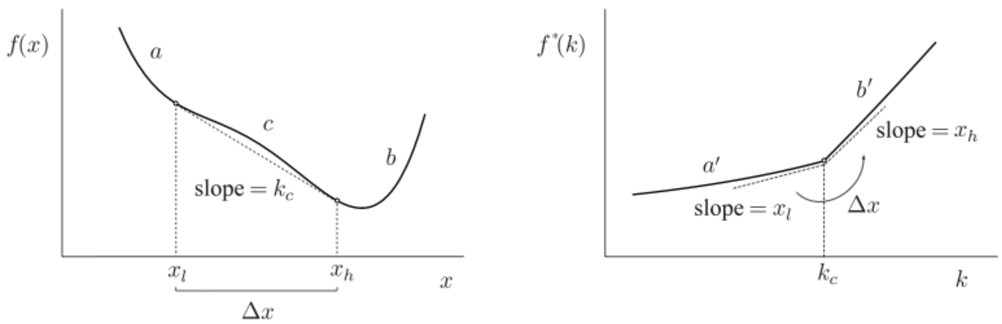
$$\partial f(y) := \left\{ k \in \mathbb{R}^n : f(x) \geq f(y) + < k, x - y > \right\}$$

- Each slope k in $\partial f(y)$ is called a subgradient. For $f(x) = |x|$, $\partial f(0) = [-1, 1]$

Principles of Duality



Function having a non-differentiable point; its LF is an affine function. (Source: Touchette 2007)



Non-convex function; its LF transform has a non-differentiable point. (Source: Touchette 2007)

Principles of Duality

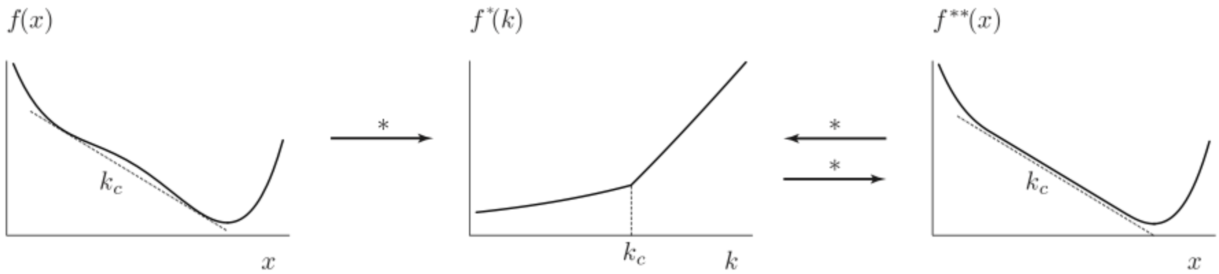
- The biconjugate

$$f^{**}(x) = \sup_{k \in \Omega_k} \Big(< k, x > - f^*(k) \Big)$$

is the largest convex l.s.c. (convex hull) function below $f(x)$

- For convex f, then $f^{**} = f$

- For non-convex f, then f^{**} is the convex envelope of f



Structure of the LF transform for non-convex functions. (Source: Touchette 2007)

Convex Problems

A class of structured convex problems

Consider the original problem:

$$\hat{u} = \arg\min_u \Big\{ \underbrace{F(Du)}_{\text{Smoothness}} + \underbrace{G(u)}_{\text{Data}} \Big\}$$

$D : \mathcal{H}_x \to \mathcal{H}_y$ a continuous linear differential operator from/to a Hilbert space

$F : \mathcal{H}_y \to \mathbb{R} \cup \{\infty\}$ and $G : \mathcal{H}_x \to \mathbb{R} \cup \{\infty\}$ are "simple" convex, proper and l.s.c. function, so they have an easy to compute proximity operator:

$$\text{prox}_{\tau F}(y) := \big(I + \tau \partial F\big)^{-1}(y) = \arg\min_x \left\{ \frac{\|x-y\|^2}{2\tau} + F(x) \right\}$$

Many problems can be cast in this framework.

Primal, Primal-Dual, and Dual Problems

Let D be a continuous linear differential operator, and F, G proper convex functions

- The primal problem corresponds to our original formulation:

$$\hat{u} = \arg\min_u \Big\{ \underbrace{F(Du)}_{\text{Smoothness}} + \underbrace{G(u)}_{\text{Data}} \Big\} \tag{1}$$

- The primal-dual problem is defined as (using $F^{**} = F$ in (1))

$$(\hat{u}, \hat{k}) = \arg\min_u \max_k \Big\{ \langle Du, k \rangle + G(u) - F^*(k) \Big\} \tag{2}$$

- The dual problem is defined as (using $\inf(h) = -\sup(-h)$ and $\langle Du, k \rangle = \langle u, D^*k \rangle$ in (2))

$$\hat{k} = \arg\max_k \Big\{ - F^*(k) - G^*(-D^*k) \Big\} \tag{3}$$

The min-max, primal-dual problem

$$(\hat{u}, \hat{k}) = \arg\min_u \max_k \Big\{ \langle Du, k \rangle + G(u) - F^*(k) \Big\}$$

- The saddle point $(\hat{u}, \hat{k})$ necessarily satisfies the Euler-Lagrange equations

- Solved by the coupled fixed point iterations (Arrow et al. 1958)

$$\begin{aligned} u^{n+1} &= \big(I + \tau \partial G\big)^{-1} \big(u^n - \tau D^* k^n\big) \\ k^{n+1} &= \big(I + \sigma \partial F^*\big)^{-1} \big(k^n + \sigma D u^{n+1}\big) \end{aligned}$$

- Faster realisations based on preconditioning (Pock, Chambolle et al. 2009–2011)

$$\begin{aligned} k^{n+1} &= \big(I + \sigma \partial F^*\big)^{-1} \big(k^n + \sigma D(u^n + \theta(u^n - u^{n-1}))\big) \\ u^{n+1} &= \big(I + \tau \partial G\big)^{-1} \big(u^n - \tau D^* k^{n+1}\big) \end{aligned}$$

with ensured convergence for $\tau\sigma \|D\|^2 < 1$ and $\theta = 1$

Summary

- Continuous energy functionals lead to systems of PDEs which need to be solved in the discrete setting.
- A continuous functional can also be directly discretised, following the computation of discrete optimality conditions.
- In both cases, we need to devise efficient solvers for the elliptic (Euler-Lagrange) and the parabolic (gradient descent) equations.
- Depending on convex (resp. non-convex) penalising functions Ψ, the energy minimiser will correspond to the global (resp. local) optimum.
- For strictly convex functionals, duality principles can be exploited to devise more efficient solvers and to handle penalisers Ψ with differentiability problems.

References

- D. M. Young: Iterative Solution of Large Linear Systems. Dover, New York, 2003.
- Y. Saad: Iterative Methods for Sparse Linear Systems. SIAM, Philadelphia, Second Edition, 2003.
- W. L. Briggs, V. E. Henson, S. F. McCormick: A Multigrid Tutorial. SIAM, Philadelphia, Second Edition, 2000.
- J. Weickert, J. Heers, C. Schnörr, K. J. Zuiderveld, O. Scherzer, H. S. Stiehl: Fast parallel algorithms for a broad class of nonlinear variational diffusion approaches. Real-Time Imaging, Vol. 7, No. 1, 31–45, Feb. 2001.
- S. Boyd, L. Vandenberghe: Convex Optimization. Cambridge University Press, 2004.
- R. T. Rockafellar: Convex analysis. Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 1997.
- K. J. Arrow, L. Hurwicz, H. Uzawa: Studies in linear and non-linear programming. Stanford Mathematical Studies in the Social Sciences, vol. II. Stanford University Press, 1958.
- T. Pock, D. Cremers, H. Bischof, A. Chambolle: An algorithm for minimizing the Mumford-Shah functional. In ICCV Proceedings, LNCS. Springer, 2009.
- A. Chambolle, T. Pock: A first-order primal-dual algorithm for convex problems with applications to imaging. Journal of Mathematical Imaging and Vision, Vol. 40, No. 1, 120–145, 2011.
- T. Pock, A. Chambolle: Diagonal preconditioning for first order primal-dual algorithms in convex optimization. International Conference on Computer Vision (ICCV 2011), pp. 1762–1769.
- D. Cremers, B. Goldlücke, T. Pock: Variational Methods for Computer Vision. ICCV Tutorial, 2011.
- H. Touchette: Legendre-Fenchel transforms in a nutshell. Technical report, School of Mathematical Sciences, Queen Mary, University of London, 2007.
- A. Handa, R. A. Newcombe, A. Angeli, A. J. Davison: Applications of Legendre-Fenchel transformation to computer vision problems. Tech.Rep. DTR11-7, Department of Computing, Imperial College, 2011.