Variational Methods in Signal and Image Processing

XU WANG
Texas A&M University
Dept. of Electrical & Computer Eng.
College Station, Texas
United States
xu.wang@tamu.edu

ERCHIN SERPEDIN
Texas A&M University
Dept. of Electrical & Computer Eng.
College Station, Texas
United States
serpedin@ece.tamu.edu

KHALID QARAQE
Texas A&M University at Qatar
Dept. of Electrical & Computer Eng.
Doha
Qatar
khalid.qaraqe@qatar.tamu.edu

Abstract: This paper provides an overview of the applications that calculus of variations found in two major fields of interest, namely signal processing and image processing. We will describe the significant role of calculus of variations in modeling and optimizing a variety of signal and image processing problems. Specifically, the key role of calculus of variations in deriving the worst-case (maximum) Cramér-Rao bound, or equivalently the minimum Fisher information, and in optimizing the pulse shaping in wireless communications will be discussed. Also, the variational methods applied in image processing, such as image denoising and deblurring, will be illustrated.

Key–Words: Calculus of variations, Fisher information, pulse shaping, image denoising, image deblurring

1 Introduction

Variational methods refer to the technique of optimizing the maximum or minimum of an integral involving unknown functions. During the last two centuries, variational methods have played an important role in addressing problems in many disciplines such as statistics, physics and particularly mechanics. Recently, the applications of variational methods have been found in some other fields such as economics and electrical engineering. The goal of this paper is to provide the readers the basic definitions, concepts and results of calculus of variations and how these results can be applied to some applications in signal and image processing.

The remainder of this paper is structured as follows. In Section 2, some variational calculus preliminary definitions and theorems will be first reviewed. In Section 3, two applications of signal processing, namely the minimum Fisher information and the pulse shaping optimization, will be formulated and solved within the framework of calculus of variations. A variational approach for denoising and deblurring an image will be discussed in Section 4. Section 5 concludes the paper.

2 Preliminaries on Calculus of Variations

In this section, some fundamental concepts and results from calculus of variations will be reviewed and used constantly throughout the rest of the paper. Most of these results are standard and therefore will be stated briefly without further details. The reader is referred to [1, 2, 3] for more details.

Definition 1. *A functional $J(f)$ of the form*

$$J(f) = \int_{x_1}^{x_2} F(x, f(x), f'(x))dx, \qquad (1)$$

is defined on the set of continuous functions $f(x)$ with continuous first-order derivatives $f'(x) = df(x)/dx$ on the interval $[x_1, x_2]$. The function $f(x)$ is assumed to satisfy the boundary conditions $f(x_1) = A, f(x_2) = B$ and $F(x, f(x), f'(x))$ is also assumed to have continuous first-order and second-order partial derivatives with respect to (w.r.t) all of their arguments.

In particular, when the functional does not depend on $f'(x)$, the equation (1) simplifies to

$$J(f) = \int_{x_1}^{x_2} F(x, f(x))dx. \qquad (2)$$

Definition 2. *Define the perturbation of $f(x)$ as*

$$\hat{f}(x) = f(x) + \epsilon\eta(x),$$

where $\eta(x)$, defined as the increment of $f(x)$, is an arbitrary continuous function with boundary conditions $\eta(x_1) = \eta(x_2) = 0$. The first variation of functional $J(f)$ is defined as

$$\delta J = \left.\frac{\partial J(\hat{f})}{\partial \epsilon}\right|_{\epsilon=0}. \qquad (3)$$

ISBN: 978-960-474-387-2

Theorem 3. *A necessary condition for the functional $J(f)$ to have an extremum (or local optimum) for a given function $f = f^*$ is that its first variation vanishes at f^*,*

$$\delta J(f^*) = 0, \tag{4}$$

for all admissible increments $\eta(x)$ defined above. This implies the Euler equation

$$F_f - \frac{d}{dx} F_{f'} = 0 \tag{5}$$

at f^*, where F_f and $F_{f'}$ denote the partial derivative $\frac{\partial F}{\partial f}$ and $\frac{\partial F}{\partial f'}$ respectively. Particularly, when the functional admits the form (2), the equation (5) reduces to

$$F_f = 0. \tag{6}$$

Theorem 4. *Suppose the functional $J(f)$ with constraints*

$$I_i(f) = \int_{x_1}^{x_2} F_i(x, f(x), f'(x))dx = l_i, \tag{7}$$

$i = 1, 2, \cdots, m$, has an extremum at f^* , then there exist constants $\alpha_1, \alpha_2, \cdots, \alpha_m$ such that f^* is also an extremum of the functional

$$J^* = \int_{x_1}^{x_2} F^* dx = \int_{x_1}^{x_2} \left(F + \sum_{i=1}^{m} \alpha_i F_i \right) dx, \tag{8}$$

and satisfies the corresponding Euler equation

$$F_f^* - \frac{d}{dx} F_{f'}^* = 0. \tag{9}$$

Note that this theorem is referred to as the Lagrange multipliers rule for isoperimetric problems.

Theorem 5. *For the functional in (1) subject to the constraints in (7), suppose (f, f') to form a convex set and $F^* = F + \sum_{i=1}^{m} F_i$ to be convex w.r.t (f, f') for each x in the interval $[x_1, x_2]$. Let f^* be an extremum satisfying the Euler equation (9), then the functional achieves the minimum at f^*.*

Proof. Assume $f(x)$ to be any arbitrary function satisfying constraints (7) and the boundary conditions, due to the fact that the convex function lies above its tangent, it turns out that

$$\begin{aligned} &F^*(x, f, f', \boldsymbol{\alpha}) - F^*(x, f^*, f'^*, \boldsymbol{\alpha}) \\ \geq &(f - f^*)F_f^* + (f' - f'^*)F_{f'}^*, \end{aligned}$$

where

$$F_f^* = \left. \frac{\partial F^*}{\partial f} \right|_{f=f^*} \text{ and } F_{f'}^* = \left. \frac{\partial F^*}{\partial f'} \right|_{f=f^*}.$$

Thus,

$$\begin{aligned} &J(f) - J(f^*) \\ =&\int F(x, f, f')dx - \int F(x, f^*, f'^*)dx \\ =&\int F(x, f, f')dx - \int F(x, f^*, f'^*)dx \\ &+ \sum_{i=1}^{m} \left[\alpha_i \left(\int F(x, f, f')dx - l_i \right) \right. \\ &\left. - \alpha_i \left(\int F(x, f^*, f'^*)dx - l_i \right) \right] \\ =&\int F^*(x, f, f', \boldsymbol{\alpha}) - F^*(x, f^*, f'^*, \boldsymbol{\alpha})dx \\ \geq&\int (f - f^*)F_f^* + (f' - f'^*)F_{f'}^* dx \\ \overset{(a)}{=}&\int (f - f^*)F_f^* dx - \int (f - f^*) \left(\frac{d}{dx} F_{f'}^* \right) dx \\ =&\int (f - f^*) \left(F_f^* - \frac{d}{dx} F_{f'}^* \right) dx \\ \overset{(b)}{=}&0, \end{aligned}$$

where (a) follows from the integration by parts and (b) is due to (9).

3 Variational Methods in Signal Processing

In signal processing, variational methods are used to model and optimize the communication systems by choosing the optimal signaling functions. In this section, we will briefly describe how variational techniques can be applied to derive the worst-case Cramér-Rao bound for unbiased estimators and to match signaling pulses optimally to channel transfer functions.

3.1 Minimum Fisher Information

In statistical signal processing, suppose θ is an unknown parameter to be determined by measurements x, which assume the probability density function (pdf) $f(x; \theta)$. Any estimator $\hat{\theta}$ of θ, which satisfies the equation $\mathrm{E}_\theta(\hat{\theta}) = \theta$, is referred to as the unbiased estimator of θ. The Cramér-Rao inequality [4] states that the variance of any unbiased estimator $\hat{\theta}$ is lower-bounded by the reciprocal of the Fisher information $I(f)$,

$$\mathrm{var}(\hat{\theta}) \geq \frac{1}{I(f)}, \tag{10}$$

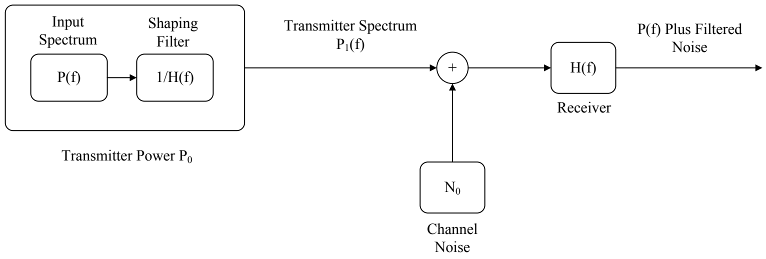

Figure 1: Communication system with pulse shaping filter

where the Fisher information $I(f)$ is defined as

$$I(f) = \int_{-\infty}^{\infty} \left(\frac{\partial \log f(x;\theta)}{\partial \theta} \right)^2 f(x;\theta)dx. \quad (11)$$

In the case of a scale or location parameter [5], the formula (11) simplifies to

$$I(f) = \int_{-\infty}^{\infty} \left(\frac{f'(x)}{f(x)} \right)^2 f(x)dx. \quad (12)$$

Now we show that if nothing is known about $f(x;\theta)$ except its second-order moment information, the Gaussian assumption is a natural choice due to the fact that the Gaussian distribution leads to the minimum Fisher information, or equivalently the maximum Cramér-Rao bound, subject to a fixed second-order moment. In the framework of calculus of variations, the problem can be formulated as

$$\begin{aligned} \min \; & \int \left(\frac{f'(x)}{f(x)} \right)^2 f(x)dx \\ \text{s.t.} \; & \int f(x)dx = 1 \\ & \int x^2 f(x)dx = \sigma^2, \end{aligned} \quad (13)$$

where the integrand is from $-\infty$ to ∞. By utilizing Theorem 4, problem (13) can be solved as

$$F^* = \left(\frac{f'(x)}{f(x)} \right)^2 f(x) + \alpha_1 f(x) + \alpha_2 x^2 f(x),$$

and

$$F_f^* - \frac{d}{dx} F_{f'}^* = \left(\frac{f'(x)}{f(x)} \right)^2 - 2\frac{f''(x)}{f(x)} + \alpha_1 + \alpha_2 x^2 = 0. \quad (14)$$

It is shown in [6] that the Gaussian distribution

$$f^*(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \quad (15)$$

is the solution of (14) with $\alpha_1 = -2/\sigma^2$ and $\alpha_2 = 1/\sigma^4$. Since the Hessian matrix of F^* w.r.t $(f(x), f'(x))$

$$\begin{bmatrix} 2f'(x)/f^3(x) & -2f'(x)/f^2(x) \\ -2f'(x)/f^2(x) & 2/f(x) \end{bmatrix}$$

is positive semidefinite, F^* is convex. Due to Theorem 5, it can be concluded that the Gaussian distribution (15) minimizes the Fisher information.

3.2 Pulse Shaping Optimization

In modern communication systems, a pulse shaping is the process of changing the waveform or spectrum of a transmitted input pulse to make the transmitted signal better suited to the characteristics of the communication channel. In this section, we will briefly discuss the variational methodology to achieve the optimal signal-to-noise ratio at the receiver side for a given total transmitted power P_0. For more details, the reader is referred to [8].

The communication system we consider is the baseband system shown in Figure 1 [8]. The transmitter has a given input power spectrum $P(f)$, an output power spectrum $P_1(f)$ and a total output power P_0. Thus, the relationship between $P(f)$ and $P_1(f)$ is given by

$$P(f) = P_1(f)H(f)H^*(f), \quad (16)$$

where $H^*(f)$ represents the complex conjugate of $H(f)$. If the spectral density is restricted to positive frequencies only, the output power spectrum $P_1(f)$ is constrained by

$$\int_0^{\infty} P_1(f)df = P_0. \quad (17)$$

Thus, the total signal power at the receiver side can be expressed as

$$P_s = \int_0^{\infty} P_1(f)H(f)H^*(f)df = \int_0^{\infty} P(f)df, \quad (18)$$

and the total noise power at the receiver side admits the form

$$P_n = \int_0^\infty N_0 H(f)H^*(f)df = \int_0^\infty N_0 \frac{P(f)}{P_1(f)} df. \tag{19}$$

Since $P(f)$ is given, maximizing the signal-to-noise ratio P_s/P_n is equivalent of minimizing the total noise power in (19). Therefore, by utilizing calculus of variations, the problem can be formulated as

$$\begin{aligned} \min & \int_0^\infty N_0 \frac{P(f)}{P_1(f)} df \\ \text{s.t.} & \int_0^\infty P_1(f)df = P_0. \end{aligned} \tag{20}$$

Based on Theorem 4, it turns out that

$$F^* = N_0 \frac{P(f)}{P_1(f)} + \alpha P_1(f),$$

and

$$\frac{\partial F^*}{\partial P_1(f)} = -N_0 \frac{P(f)}{[P_1(f)]^2} + \alpha = 0,$$

which yields the solution

$$P_1^*(f) = \sqrt{\frac{N_0 P(f)}{\alpha 1}}. \tag{21}$$

Note that $P_1^*(f)$ in (21) is the optimal solution of (20) since F^* is convex w.r.t $P_1(f)$. Plugging (21) into the constraint (17) leads to

$$\alpha = \frac{N_0}{P_0^2}\left(\int_0^\infty \sqrt{P(f)}df\right)^2,$$

and the optimal signal-to-noise ratio admits the form

$$\frac{P_s}{P_n} = \frac{P_0}{N_0}\frac{\int_0^\infty P(f)df}{\left(\int_0^\infty \sqrt{P(f)}df\right)^2}. \tag{22}$$

Therefore, the optimal signal-to-noise ratio for any prescribed signal spectral shape $P(f)$ is given by (22) and the shaping filter $H(f)$ can be designed by using the relationships in (16) and (21).

4 Variational Methods in Image Processing

Traditionally, image processing has been investigated via spectral and Fourier methods in the frequency domain. In the last two decades, variational methods and partial differential equation (PDE) methods have drawn great attention to address a variety of image processing problems including image segmentation, image registration, image denoising and image deblurring. In this section, we will focus our attention on applications of variational techniques in image denoising and image deblurring.

Figure 2: Original image

4.1 Image Denoising

In the process of acquiring, storing and transmitting an image, the measurements are always perturbed by noise. An example of a noisy image is given in Figure 3 compared to the original image in Figure 2

Mathematically, the model can be expressed as

$$u_0 = u + n, \tag{23}$$

where the observed image u_0 includes the original image u and the additive noise n. The method that removes the noise from the observed image is called image denoising. The simplest and the best investigated way for denoising noise is to apply a linear filter such as the Gaussian smoothing filter and the Wiener filter [9]. Recently, variational methods have been introduced to address the image denosing problem. Instead of considering an image as a sampled and quantized matrix, we define an image to be a continuous real-valued function $u : \Omega \to \mathbb{R}$, where $\Omega = \{(x, y)|a < x < b, c < y < d\}$. When considering the signal model (23), the variational method involves determining the unknown function u that minimizes the objective functional:

$$\iint_\Omega F(u)dxdy, \tag{24}$$

subject to the image data constraint

$$\iint_\Omega (u - u_0)^2 dxdy = \sigma^2. \tag{25}$$

Extending the Lagrange multipliers rule in Theorem 4 to the two-dimensional case [1], the problem resumes to

$$\min \iint_\Omega F(u) + \frac{\alpha}{2}(u - u_0)^2 dxdy, \tag{26}$$

ISBN: 978-960-474-387-2

Figure 3: Noisy image

where α is an nonnegative coefficient parameter so that the constraint (25) is satisfied. In practice, α is often estimated or chosen as a prior information [10]. The functional in (26) is often referred to as the energy functional in image processing, where $(u - u_0)^2$ represents the data term or similarity term, and $F(u)$ represents the smoothing term or regularization term [11].

Consequently, a critical issue is to determine the smoothing term $F(u)$. A special aspect of image processing within the framework of variational methods is to consider the edges at which discontinuities occur. This can be done by using the total variation, which is defined as the average magnitude of the gradient of u over the region Ω [12, 13]:

$$TV(u) = \iint_\Omega |\nabla u| dxdy = \iint_\Omega \sqrt{u_x^2 + u_y^2} dxdy. \tag{27}$$

Another well-known functional that is also designed for allowing abrupt changes within images is referred to as the Dirichlet variation [10]:

$$D(u) = \frac{1}{2} \iint_\Omega |\nabla u|^2 dxdy. \tag{28}$$

Consider the problem (26) under the assumption that the smoothing term $F(u)$ is modeled as the total variation defined in (27) and define

$$J(u) = \iint_\Omega |\nabla u| + \frac{\alpha}{2}(u - u_0)^2 dxdy, \tag{29}$$

then the problem can be formulated by

$$\hat{u} = \arg\min J(u). \tag{30}$$

It can be seen that $J(u)$ admits the general functional form

$$J(u) = \iint_\Omega F^*(x, y, u, u_x, u_y) dxdy, \tag{31}$$

which represents a two-dimensional extension of Definition 1. It is shown in [1] that Theorems 3 and 4 can be generalized to attain the corresponding Euler equation (31) as follows

$$F_u^* - \frac{\partial}{\partial x}(F_{u_x}^*) - \frac{\partial}{\partial y}(F_{u_y}^*) = 0. \tag{32}$$

As a result, the corresponding partial differential Euler equation of (29) is given by

$$\nabla \left(\frac{\nabla u}{|\nabla u|} \right) - \alpha(u - u_0) = 0, \tag{33}$$

or equivalently

$$\frac{\partial}{\partial x}\left(\frac{u_x}{\sqrt{u_x^2 + u_y^2}} \right) + \frac{\partial}{\partial y}\left(\frac{u_y}{\sqrt{u_x^2 + u_y^2}} \right) - \alpha(u - u_0) = 0. \tag{34}$$

To solve the PDE in (33), iterative methods such as gradient descent [13], or fixed point method [14] may be applied.

4.2 Image Deblurring

Image blurring, which is referred to as the process of recovering the original image from the blurred image, usually occurs because of the movement of the imaging device during shooting or choosing a fast enough shutter speed to freeze the action under the light conditions. An example of image blurring is shown in Figure 4 as a comparison to the original image in Figure 2.

Image blurring is typically modeled with a linear and shift-invariant operator K defined as the convolution of the image [11]:

$$K(u) = k(x, y) * u(x, y),$$

where $k(x, y)$ denotes the convolution kernel and $K(u)$ represents the convolution of the image. Under the assumption that the total variation (27) is used as the smoothing term, the energy functional is given by

$$J(u) = \iint_\Omega |\nabla u| + \frac{\alpha}{2}(K(u) - u_0)^2 dxdy. \tag{35}$$

Similarly to the image denoising case, the image deblurring model (35) is computed by the Euler equation (32), which resumes to:

$$\nabla \left(\frac{\nabla u}{|\nabla u|} \right) - \alpha K^*(K(u) - u_0) = 0, \tag{36}$$

Figure 4: Blurred image

or equivalently

$$\frac{\partial}{\partial x}\left(\frac{u_x}{\sqrt{u_x^2+u_y^2}}\right)+\frac{\partial}{\partial y}\left(\frac{u_y}{\sqrt{u_x^2+u_y^2}}\right)-$$
$$\alpha K^*(K(u)-u_0)=0,$$

where K^* represents the adjoint operator.

If we define the notation $|x|_a$ as $|x|_a = \sqrt{x^2+a^2}$ for a fixed positive parameter a, the nonlinear degenerate elliptic equation (36) is often regularized to [15]:

$$\nabla\left(\frac{\nabla u}{|\nabla u|_a}\right)-\alpha K^*(K(u)-u_0)=0,$$

which can be computationally solved by an iterative approach referred to as the lagged-diffusivity technique [16, 17].

5 Conclusion

In this paper, several applications of variational techniques were discussed mainly in signal and image processing area. Rather than illustrate mathematical details of calculus of variations and the aforementioned applications, our goal is to provide the reader some insight into how signal and image processing problems can be formulated and tackled within the framework of calculus of variations.

Acknowledgements: This paper was supported by NPRP grant 4-1293-2-513 from the Qatar National Research Fund (a member of Qatar Foundation). The statements made herein are solely the responsibility of the authors.

References:

[1] I. Gelfand and S. Fomin, *Calculus of Variations,* Dover, New York 1991.

[2] B. Brunt, *The Calculus of Variations,* Springer, New York 2004.

[3] G. Korn and T. Korn, *Mathematical Handbook for Scientists and Engineers,* Dover, New York 2000.

[4] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory,* Prentice Hall, Upper Saddle River 2008.

[5] K. Kim and G. Shevlyakov, Why Gaussianity?, *IEEE Signal Processing Magzine* 25(2), 2008, pp. 102–112.

[6] J. Bercher and C. Vignat, On minimum Fisher information distributions with restricted support and fixed variance, *Information Science* 179, 2009, pp. 3832–3842.

[7] E. Uhrmann Klingen, Minimal Fisher information distributions with compact-supports, *The Indian Journal of Statistics* 57(3), 1995, pp. 360–374.

[8] J. Pierce and E. Posner, *Introduction to Communication Science and Systems,* Springer, New York 1980.

[9] R. Gonzalez and R. Woods, *Digital Image Processing,* Prentice Hall, Upper Saddle River 2007.

[10] A. Hamza, H. Krim and G. Unal, Unifying probabilistic and variational estimation, *IEEE Signal Processing Magzine* 19(5), 2002, pp. 37–47.

[11] K. Cassel, *Variational Methods with Applications in Science and Engineering,* Cambridge University Press, Cambridge 2013.

[12] W. Allard, Total variation regularization for image denoising I: Geometric theory, *SIAM J. Math. Anal.* 39(4), 2007, pp. 1150–1190.

[13] L. Rudin, S. Osher and E. Fatemi, Nonlinear total variation based noise removal algorithms, *Physica D* 60, 1992, pp. 259–268.

[14] K. Ito and K. Kunisch, Restoration of edge-flat-grey scale image, *Inverse Problems* 16(4), 2000, pp. 909–928.

[15] T. Chan and J. Shen, Theory and computation of variational image deblurring, *Math. Comput. Imaging Sci. Inf. Process,* 2007, pp. 93–131.

[16] A. Chambolle and P. Lions, Image recovery via total variational minimization and related problems, *Numer. Math.* 76, 1997, pp. 167-188.

[17] T. Chan and J. Shen, *Image Processing and Analysis: Variational, PDE, Wavelet, and Stochastic Methods,* SIAM Publisher, Philadelphia 2005.