<u>Variational Methods</u>

Contents

Math Intro

Noise Modeling

Noise Modelling

We write $f(i,j) = s(i,j) + n(i,j)$, where s is true signal, n is noise, and f is the acquired signal (noisy).

Noise can be modeled by a histogram or a PDF which is superimposed on the PDF of the original image s.

$\ast$ Def. **0.0.1 Salt And Pepper Noise:** n takes only two values - 0 (black, pepper) or 255 (white, salt).

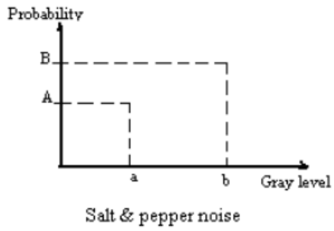

Fig. 10.1 Probability density function for the salt & pepper noise model.

$$PDF_{salt\&pepper} = \begin{cases} A & for\ g = a\ ("pepper") \\ B & for\ g = b\ ("salt") \end{cases} \qquad (10.2)$$

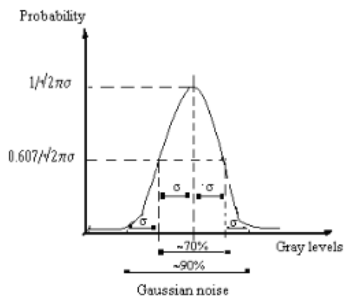

Fig. 10.2 Probability density function for the Gaussian noise model

$$PDF_{Gaussian} = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(g-\mu)^2}{2\sigma^2}} \qquad (10.3)$$

where:

g = gray level;
μ = mean;
σ = standard deviation;

Figure 1: Gaussian noise

∗ Def. **0.0.2 Gaussian Noise:**

∗ Def. **0.0.3 White Noise:** A random vector is said to be a white noise vector if its components each have a PDF with zero mean and finite variance, **and are statistically independent**. That is, their joint PDF must be the product of the distributions of the individual components.

If, in addition to being independent, every variable in the vector w has a normal distribution with zero mean and the same variance σ^2, this vector is said to be a Gaussian white noise vector. The joint distribution of w is the multivariate normal distribution.

1 Week 1

1.1 Energy and Optimization

∗ Def. **1.1.1 Acquisition Model:**

$$\boxed{f = g * H + n}$$

where f is observed image (signal), g is clean image, n is AWGN, and H is known/estimated blur kernel.

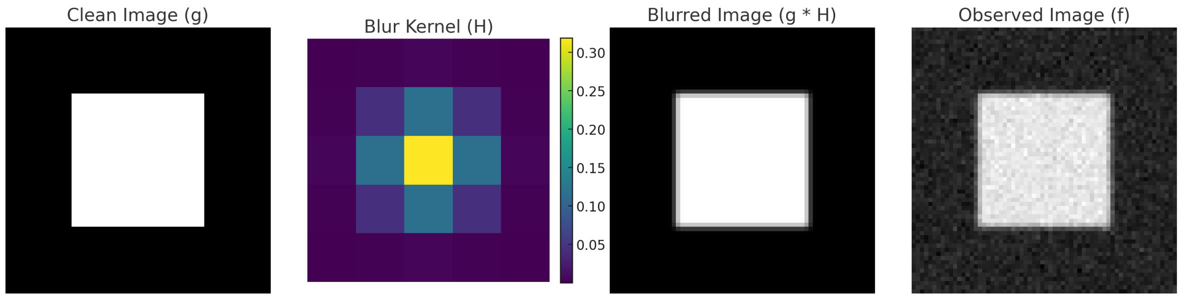

Figure 2: Acquisition Process

∗ Def. **1.1.2 Total Variation Energy:** Given a signal u, we define its total variation by:

$$J_{TV}(u) \stackrel{\text{def}}{=} \int_{\Omega} |\nabla u(x)| \, dx$$

∗ Def. **1.1.3 Fidelity Energy:** For two signals u, v Fidelity measures how similar are they. For example, we can use the $L2$ norm:

$$E_{\text{Fidelity}} = \|u - v\|_2^2$$

@ Example Optimization Formulation of the Acquisition Process: Let $f = g * H + n$ as defined above. We can define the problem of finding g as :

$$\text{Solution } u, \quad u = \arg\min_u E_{\text{smoothness}}(u) + \lambda E_{\text{Fidelity}}(u, f) = J_{TV}(u) + \|f - u * H\|_2^2$$

1.2 Calculus of Variations

In finite-dimensional calculus, the derivative of a function provides a linear approximation that describes how the function changes in response to small changes in its input. When we extend our study to infinite-dimensional spaces, such as function spaces in functional analysis, we need a generalized notion of derivatives to capture how functions behave in these settings. This leads us to the concepts of Gâteaux and Fréchet derivatives, which serve as tools to analyze and approximate functions between infinite-dimensional normed vector spaces (Banach spaces).

What We Expect of Such Derivatives:

- Linearity: The derivative should be a linear operator that approximates the change in the function.
- Continuity: For the approximation to be meaningful, especially in the Fréchet sense, the linear operator should be continuous.
- Directional Sensitivity: The Gâteaux derivative captures the rate of change of the function in a specific direction, similar to the directional derivative in finite dimensions.
- Uniform Approximation: The Fréchet derivative provides a uniform approximation over all directions, akin to the total derivative.

$\ast$ Def. **1.2.1 Gâteaux Derivative:** Let X be a Banach space, $u, v \in X$, and $E : X \to \mathbb{R}$. The Gâteaux Derivative of E at $u \in X$ is defined as:

$$\delta E(u; v) = E'(u; v) \overset{\text{def}}{=} \lim_{\lambda \to 0^+} \frac{E(u + \lambda v) - E(u)}{\lambda} = \frac{dE(u + \lambda v)}{d\lambda}\Big|_{\lambda=0}$$

We call v a **variation**, we call λv a **perturbation**. We assume zero contribution on the boundary, i.e. $v|_\Omega = 0$. We also assume that E is Gâteaux differentiable, that is, for all v we reach the same limit. We call $\delta E(u; v)$ the **first variation** or **Gâteaux variation** of E at $u \in X$.

The existence of such the Gâteaux derivative presupposes that:

- $E(u)$ is defined
- $E(u + \lambda v)$ is defined for all sufficiently small λ.

Remark (): In the expression $\frac{dE(u+\lambda v)}{d\lambda}$, we are really looking at the function $\lambda \to E(u + \lambda v)$, and differentiating it with respect to λ. The expression $\frac{dE(u+\lambda v)}{d\lambda}\Big|_{\lambda=0}$ assumes that $\frac{dE(u+\lambda v)}{d\lambda}$ is defined at $\lambda = 0$.

\# Thm. **1.2.2 Uniqueness of the Gâteaux Derivative:**

$\ast$ Def. **1.2.3 Gâteaux Differentiable:** We say a function $E : X \to Y$ is Gâteaux differentiable at $u \in X$, if there is a bounded and linear operator $D_E(u) : X \to Y$ such that :

$$\lim_{\lambda \to 0} \frac{E(u + \lambda v) - E(u)}{\lambda} = D_E(u)(v)$$

for every $v \in X$. We call the operator $D_E(u)$ the **Gâteaux derivative** of E at u. It gives the instantaneous rate of change E at u in the direction of v.

@ **Example:** Let X be $C^1[a, b]$, and let $E : C^1[a, b] \to \mathbb{R}$ be $E(u) = \int_a^b |u(t)|^2\, dt$. Then, for each $u \in C^1[a, b]$ we have that the integral is finite, and hence E is defined. Let $v \in C^1[a, b]$ be an arbitrary direction. Then:

$$\frac{E(u + \lambda v) - E(v)}{\lambda} = \frac{1}{\lambda} \int_a^b (u(x) + \lambda v(x))^2 - u(x)^2\, dx = 2 \int_a^b u(x)v(x)\, dx + \lambda \int_a^b (v(x))^2\, dx$$

Letting $\lambda \to 0$, we get:

$$\frac{dE(u + \lambda v)}{d\lambda}\Big|_{\lambda=0} = \delta E(u; v) = 2 \int_a^b \big(u(x)v(x)\big)\, dx$$

Or equivalently, define $f(\lambda) = E(u + \lambda v) = \int_a^b |(u + \lambda v)(x)|\, dx$. Then $f'(\lambda) = 2 \int_a^b (uv)(x)\, dx + \lambda \int_a^b v(x)^2\, dx$, and $\delta E(u; v) = f'(\lambda)\Big|_{\lambda=0} = f'(0) = 2 \int_a^b (uv)(x)\, dx$.

Since this exists for all $v \in C^1[a, b]$, we say that E is Gâteaux differentiable at $u \in C^1[a, b]$, and we have its *Gâteaux derivative*:

$$D_E(u)(v) = 2 \int_a^b \big(u(x)v(x)\big)\, dx$$

@ **Example Non Gâteaux differentiable:** Let $E : C^1[0, 1] \to \mathbb{R}$ be the functional defined, at $u \in C^1[0, 1]$ by:

$$E(u) := \int_0^1 |u(x)|\, dx$$

Then, at "point" $u(x) := 0$, and for direction $v(x) := x$, we have :

$$E(u + \lambda v) = \int_0^1 |u(x) + \lambda v(x)|\, dx = \int_0^1 |0 + \lambda x|\, dx = |\lambda| \int_0^1 |x|\, dx = \frac{|\lambda|}{2}$$

And:

$$E(u + \lambda v) - E(u) = \frac{|\lambda|}{2}$$

i.e.:

$$\frac{E(u + \lambda v) - E(v)}{\lambda} = \frac{1}{\lambda} \frac{|\lambda|}{2} = \begin{cases} 1/2, & \lambda > 0 \\ -1/2, & \lambda < 0 \\ \text{undefined}, & \lambda = 0 \end{cases}$$

Therefore: $\delta E(u;v)$ does not exist, since
$$\lim_{\lambda\to 0}\frac{E(u+\lambda v)-E(u)}{\lambda}$$
does not exist, particularly because it has different values when approaching from different sides.

Equivalently, we can define
$$f(\lambda)=E(u+\lambda v)=\int_0^1 |u+\lambda v|(x)\,dx=|\lambda|\int_0^1 |x|\,dx=|\lambda|\frac{1}{2}.$$
Differentiating with respect to λ, we get:
$$f'(\lambda)=\begin{cases}1/2, & \lambda>0\\ -1/2, & \lambda<0\\ \text{undefined}, & \lambda=0\end{cases}$$
Therefore we have that $\delta E(u;v)=f'(0)$ is undefined.

$\ast$ Def. **1.2.4 Fréchet Derivative:** Let X, Y be normed vector spaces, and $f: X\to Y$. The Fréchet derivative of f at a point $x\in X$ is a bounded linear operator $A: X\to Y$ such that:
$$\lim_{\|h\|\to 0}\frac{\|f(x+h)-f(x)-Ah\|_Y}{\|h\|_X}=0$$

1.3 Functional Derivative

$\ast$ Def. **1.3.1 Functional Differential:** Let X be a Banach space, and E a functional defined on X. The **differential** of E at point $u\in X$, is the linear functional $\delta E[u,\cdot]$ on X defined by the condition that for all directions $v\in X$:
$$E[u+v]-E[u]=\delta E[u;v]+\lambda\|v\|$$
where λ is a small number that depends on $\|v\|$ in such a way that $\lambda\to 0$ as $\|v\|\to 0$. The means that $\delta E[u,\cdot]$ is the Fréchet Derivative of E at u.

However, this notion of functional differential is so strong it may not exist, and in those cases a weaker notion, like the Gâteaux derivative is preferred. Thus, the functional differential is defined as the directional derivative:
$$\delta E[u,v]:=\lim_{\lambda\to 0}\frac{E[u+\lambda v]-E[u]}{\lambda}=\left[\frac{d}{d\lambda}E[u+\lambda v]\right]_{\lambda=0}$$

$\ast$ Def. **1.3.2 Functional Derivative:** The **functional derivative** $\frac{\delta E}{\delta u}$ of a functional E at a point u measures the sensitivity of E to small variations in the function u.

Assume the domain of the functional E is the space of differentiable functions u defined on some space Ω, and E is of the form:
$$E[u]=\int_\Omega L\big(x,u(x),Du(x)\big)\,dx$$
if this is the case, and, moreover, the functional differential $\delta E[u,v]$ can be written as the integral of v times another function, denoted $\delta E/\delta u$:
$$\delta E[u,v]=\int_\Omega \frac{\delta E}{\delta u}(x)v(x)\,dx$$
then this function $\delta E/\delta u$ is called the functional derivative of E at u.

Remark (functional derivative and first-order approximation): The functional derivative $\frac{\delta E}{\delta u}$ is defined such that for a small perturbation $v(x)$, we have:
$$E[u+v]-E[u]=\int\frac{\delta E}{\delta u}(x)v(x)\,dx+o(\|v\|)$$
where $o(\|v\|)$ represent terms of higher order than linear in v i.e.
$$E[u+v]=E[u]+\int\frac{\delta E}{\delta u}(x)v(x)\,dx+o(\|v\|)$$
Therefore we see that the first-order approximation of E at $u+v$ near u, is the integral of the functional derivative times the delta v.

@ Example: In example 1.2, we've defined $E : C^1[a,b] \to \mathbb{R}$ by

$$E(u) = \int_a^b |u(t)|^2 \, dt$$

and saw that:

$$\delta E(u; v) = \int_a^b 2u(x)v(x) \, dx$$

Therefore,

$$\frac{\delta E}{\delta u}(x) = 2u(x)$$

@ Example: Let

$$L[P(x)] = \int P(x) \log P(x)$$

Recall that near a point x, we have that $\log(x + \Delta x) = \log x + \frac{\Delta x}{x} + o(\Delta x)$.

Let δP denote a small perturbation of P. Therefore, $\log(P + \delta P) \approx \log P + \frac{\delta P}{P}$ (valid for $\delta P/P << 1$). Therefore

$$\begin{aligned} L[P + \delta P] &= \int \Big(P + \delta P\Big) \log \big(P + \delta P\big) \, dx \\ &\approx \int (P + \delta P)(\log P + \frac{\delta P}{P}) \, dx \\ &= \int P \log P + \delta P + \delta P \log P + \frac{(\delta P)^2}{P} \, dx \end{aligned}$$

neglecting $\frac{(\delta P)^2}{P}$, we get:

$$L[P + \delta P] - L[P] = \int (1 + \log P) \delta P \, dx$$

Which is exactly of the form of the functional derivative. Hence:

$$\frac{\delta L}{\delta P} = 1 + \log P(x)$$

2 Perona-Malik

Thm. **2.0.1 Finite Difference and Flow:** Let u be a 1D array.

- Forward-Difference at i measures **inflow** from right: Define forward difference at i by:

  $$u[i+1] - u[i]$$

  When $u[i+1] > u[i]$ there will be flow from $i+1$ into i. Therefore the FD at i measures the (signed) inflow from $i+1$ to i.

- Negative Forward-Difference at i measures **outflow** from i to right The expression:

  $$u[i] - u[i+1]$$

  when $u[i] > u[i+1]$, there will be flow from i to $i+1$, hence this measures **outflow** from i to $i+1$.

  Specifically, when at $i-1$, the expression:

  $$u[i-1] - u[i]$$

  measures **outflow** from $(i-1)$ to i, which is **inflow** into i from $i-1$.

Thm. **2.0.2 Perona-Malik:** Fick's Law tells us that

$$\boldsymbol{J} = -c \boldsymbol{\nabla} u \tag{1}$$

Also the continuity equation is:

$$\frac{\partial u}{\partial t} = -\operatorname{div} \boldsymbol{J} \tag{2}$$

Hence $\boldsymbol{J}$ measures the flux, i.e. the net **outflow** from an infinitesimal volume, like the outflow from pixel (i,j)

Combining Fick's Law with the continuity equation, we get the Perona-Malik equation:

$$\frac{\partial u}{\partial t} = \mathrm{div}(c\boldsymbol{\nabla} u)$$

I.e. the amount of change of u is equal to $\mathrm{div}(c\boldsymbol{\nabla} u)$. Hence $\mathrm{div}(c\boldsymbol{\nabla} u)_{i,j}$ measures the net **inflow** into (i,j) pixel.

Focusing on the x-axis, we wish to measure the net **inflow** into (i,j) pixel along the x-axis, which is $\mathrm{div}_x(c\boldsymbol{\nabla} u)_{i,j}$

At (i,j) we have:

$$\begin{aligned}\mathrm{div}_x(c\boldsymbol{\nabla} u)_{i,j} = \text{net inflow into } (i,j) &= \Big(\text{inflow into } (i,j) \text{ from } (i,j+1)\Big) + \Big(\text{inflow into } (i,j) \text{ from } (i,j-1)\Big) \\ &\approx \Big(u[i,j+1] - u[i,j]\Big) \qquad\qquad\quad + \Big(u[i,j-1] - u[i,j]\Big)\end{aligned}$$