图像复原优化算法

张 彬 于欣妍 朱永贵 著

Optimal algorithm for image restoration

国防工業出版社
National Defense Industry Press

图像复原优化算法

张　彬　于欣妍　朱永贵　著

国防工业出版社
·北京·

内容简介

本书以作者为研究生开设的“反问题的计算方法”课程讲义为基础，对图像复原算法的原理和基础进行了较为深入的讨论，力求使读者从原理上掌握相关算法，并能用来解决实际问题。全书共分8章，主要包括：基于正则化的图像复原算法、Bregman分裂算法及其应用、基于偏微分方程的图像复原算法、变指数函数空间在图像复原和增强中的应用、深度学习在图像去雨、单目避障系统和图像复原中的应用。

本书可作为理工科各专业本科生和研究生学习图像复原的教材，也适合对图像复原算法感兴趣的研究人员和工程技术人员阅读。

图书在版编目（CIP）数据

图像复原优化算法/张彬，于欣妍，朱永贵著. —北京：国防工业出版社，2019.8

ISBN 978-7-118-11924-4

Ⅰ. ①图…　Ⅱ. ①张…　②于…　③朱…　Ⅲ. ①图像恢复－最优化算法

Ⅳ. ①TN911.73

中国版本图书馆CIP数据核字（2019）第140417号

※

国防工業出版社出版发行

（北京市海淀区紫竹院南路23号　邮政编码 100048）

国防工业出版社印刷厂印刷

新华书店经售

*

开本 787×1092　1/16　**印张** 17¼　**字数** 425千字

2019年8月第1版第1次印刷　**印数** 1—800册　**定价** 168.00元

国防书店：（010）68428422　　发行邮购：（010）68414474

发行传真：（010）68411535　　发行业务：（010）68472764

前　言

在数字图像处理的过程中，由于成像系统、传输介质、记录设备和处理方法的不完善，图像在形成、传输、记录和处理过程中，常常受到摄像机聚焦不佳、传感器噪声、光学系统的像差、光学成像衍射、成像系统的非线性畸变、物体与摄像机之间的相对移动、随机大气湍流等因素的影响，造成不同程度的图像质量下降，无法完全反映场景的真实内容，这种现象被称为图像退化。

图像复原是一种改善图像质量的技术，从分析图像退化的机理入手，直接针对产生退化的原因，根据图像退化的先验知识进行分析，在此基础上建立数学模型，通过一定的求解策略对模型进行求解，最终得出原始清晰图像的最佳估计。早期的图像复原是利用光学的方法对失真的观测图像进行校正，而数字图像复原技术最早则是从对天文观测图像的后期处理中逐步发展起来的。其中一个成功例子是 NASA 的喷气推进实验室在 1964 年用计算机处理有关月球的照片。照片是空间飞行器用电视摄像机拍摄的，图像的复原包括消除干扰和噪声、校正几何失真和对比度损失以及反卷积。另一个典型的例子是对肯尼迪遇刺事件现场照片的处理。由于事发突然，照片是在相机与拍摄对象发生相对移动过程中拍摄的，对其图像复原的主要目的就是消除移动造成的失真。

在图像退化的过程中，常会有噪声的干扰，使得求解过程变得很不稳定，为此引入正则化方法，将病态问题转化为良态问题，使求解过程变得稳定。目前，图像复原算法主要分为非盲图像复原算法和图像盲复原算法两大类。

基于偏微分方程的图像去噪是图像复原的一个重要研究方向，这方面的研究最早可以追溯到 1965 年 Gabor 及 Jain 等人的工作，但是这个领域中起奠基性作用的要归功于 Koend- erink 和 Witkin，他们创造性地提出了尺度空间概念，提出高斯滤波处理等效于热传导方程的各向同性扩散。1986 年，Hummel 指出经过多尺度滤波后得到的图像，可以看作各向同性热扩散方程的解，并且他还注意到热扩散方程并不是唯一的能够产生尺度空间的偏微分方程，满足极值原理的进化方程同样可以产生尺度空间。1987 年，Kass 等人提出了非常著名的主动轮廓模型（Active Contour Model，ACM），该模型建立了一种以图像边缘的内力/外力相约束的、能够表征图像区域轮廓的能量泛函，它以内力约束曲线的形貌，以外力引导扩散的行为，最终使得曲线逼近于最为理想的特征。1988 年，Osher 等人提出了具有几何约束的偏微分方程——曲率驱动（Curvature Driven）的扩散方程，该方程为各向异性扩散模型，即图像灰度值仅沿其水平集方向（图像梯度的正交方向）扩散，这样可以保护图像的边缘特征。

20 世纪 90 年代初期图像复原问题又发展出两大重要方法：一个是 1990 年 Perona 和 Malik 在基于热传导方程的基础上提出的各向异性扩散方程，即用一个可以保持边缘的有选择性的扩散来替换 Gaussian 扩散从而保留图像边缘的信息，人们一般将他们建立的这个模型称为 P-M 模型，P-M 方程具有各向异性扩散的特点，能够在去除噪声的同时保护边缘。另一个是 1992 年 Rudin、Osher 和 Fatemi 将 Tikhonov 正则化模型中 L_2 范数约束条件改为 L_1 范数，从而建立

了图像复原问题的变分模型，该模型能够在 L_1 范数空间中描述图像灰度的跳变，即图像梯度信息。人们一般将他们建立的这个模型称为 ROF 模型。ROF 模型在 L_1 范数空间中的正则化过程实际就是各向异性扩散过程，所以该模型也能够在去除噪声的同时保护图像边缘。这两种方法各自形成了十分丰富的研究体系。尽管相对于传统的滤波器和各向同性的热方程，P-M 模型和 ROF 模型都具有增强边缘、保持边缘细节的优点，但也同样存在现今熟知的阶梯效应。

近年来，在图像复原模型的求解方面，Bregman 迭代因其简单、稳定、速度快和效率高而受到越来越多的关注。Bregman 迭代是一系列以 Bregman 距离为基础的方法的总称，包括经典的 Bregman 方法、线性 Bregman 方法和分裂 Bregman 方法等。Bregman 迭代方法的基本思想是将优化问题分解为等价的非约束优化子问题，其中某些子问题的目标函数由 Bregman 距离定义。该方法最早由 Osher 等人引入图像复原领域，用以改进传统方法对 TV 模型的处理。为进一步提高经典 Bregman 方法的性能，Goldstein 等人提出了适用性更好的 L_1 正则化问题的分裂 Bregman 方法。它可看作 Bregman 方法和算子分裂（Operator Splitting）相结合的一种方法，相比传统的罚函数法，该算法具有收敛速度更快，更稳定，参数在迭代过程保持不变的优点。

本书共 8 章，第 1 章介绍相关算法的数学基础，第 2 章讲述的是分块循环矩阵和分块 Toeplitz 矩阵与二维离散傅里叶变换之间的关系，第 3 章讲述的是在不同的边界条件下利用二维离散傅里叶变换和共轭梯度法对基于吉洪诺夫正则化的图像复原算法求解问题，第 4 章讲述的是基于全变差正则化的图像复原算法，第 5 章讲述的是 Bregman 分裂算法及其在图像盲复原中的应用，第 6 章讲述的是基于偏微分方程的图像去噪问题，第 7 章讲述的是变指数函数空间在图像盲复原和图像增强中的应用，第 8 章讲述的是深度学习在图像处理中的有关应用。

本书在出版过程中得到国家自然科学基金项目基于压缩感知的核磁共振成像问题驱动的应用数学研究（11571325）、中国传媒大学国家自然科学基金项目配套经费（PT-61071148）、中国传媒大学中央高校基本科研业务费专项资金（CUC2019A002, CUC2019B021）和中国传媒大学青年理工科规划项目（3132018XNG1831）的资助，在此表示感谢。

在本书编写过程中，豆泽阳、相鹏、华敏杰在图像盲复原和深度学习方面提供了宝贵的材料和成果，在编程过程中得到了豆泽阳、陈涛的有益建议和帮助，北京理工大学许廷发教授、中北大学杨风暴教授审阅了书稿，并提出了宝贵的修改意见，在此一并表示感谢！本书所参考的文献已在书后列出，在此向这些文献的作者表示诚挚的感谢。限于作者水平，书中难免存在不妥和错误之处，殷切期望各位专家、学者批评指正。

张 彬　于欣妍　朱永贵
2018 年 10 月于中国传媒大学
数据科学与智能媒体学院

目 录

第 1 章　相关数学基础

图像复原问题主要涉及图像去模糊和去噪声两个方面。由光的衍射、散焦、大气湍流、相机与场景的相对位移、光学仪器频谱增益的约束等造成的图像模糊一般发生于成像过程。噪声的出现则始于图像传输记录过程，在此过程中光照程度、传感元器件自身的质量及传输过程中传输信道受到的干扰等原因会产生大量噪声，各种在图像记录存储过程中发生的测量误差、量化计算误差也是造成图像中掺杂随机噪声的原因。从退化图像寻找清晰图像的过程是图像退化的一个逆过程，由此导致图像复原问题是一个典型的反问题。由于原问题形成的复杂性，导致反问题通常具有不适定性，即不满足解的存在性、唯一以及稳定性。故克服反问题的不适定性是解决反问题的关键。自 Tikihonov 提出正则化方法以来，反问题获得了一种修正其病态性的有效方法。 图像复原问题的正则化方法是在空间域上以符合图像特征的先验信息作为约束条件，通过寻找与模糊图像最近似的清晰图像从而获得复原问题的解。图像的先验信息在最优化模型中是以正则项的形式出现，它直接影响到复原图像的质量。基于变分法和最优化方法为理论基础的正则化方法凭借其模型简洁、适定性分析相对容易及有较多数值求解技术支持而受到数字图像处理领域学者的普遍关注。

在进行图像复原时，往往归结为求解无约束优化问题 $\min\limits_{u\in M} J(u)$，对这类优化问题的存在性有必要进行深入的讨论，经典的 Weierstrass 定理要求泛函 $J(u)$ 在非空紧集 M 上连续就能保证此最优值的存在，但在无限维空间上，单位闭球不再是一个紧集，因此需要引入序列的弱收敛概念，建立更一般的最优值存在定理。在最优解存在的前提下，对这些正则化问题，通过变分的方法将其转化为满足一定边界条件的 Euler-Lagrange 方程进行数值求解，即得复原图像。

1.1　算子方程的病态性

定义 1.1.1　设有算子 $K: H_1 \to H_2$，算子方程 $K(f)=g$ 称为是适定的，则满足如下条件：

（1）对每一个 $g\in H_2$，存在 $f\in H_1$，使得 $K(f)=g$ 成立；

（2）算子方程 $K(f)=g$ 的解是唯一的；

（3）方程的解对于 g 的扰动（误差）是稳定的，即如有 $K(f_*)=g_*$，$K(f)=g$，那么当 $g\to g_*$ 时有 $f\to f_*$。

如果方程 $K(f)=g$ 不满足上述三条中的任一条，则称该方程是病态的。

例 1.1.1　考虑 $l^2(R)$ 上的对角算子，对 $\boldsymbol{f}=(f_1,f_2,\cdots)\in l^2(R)$ 有

$$\boldsymbol{D}(\boldsymbol{f})=\left(f_1,\frac{1}{2}f_2,\cdots,\frac{1}{n}f_n\cdots\right)$$

显然这是一个单射，但不是满射。如取 $\boldsymbol{g}=\left(1,\frac{1}{2},\frac{1}{3},\cdots\right)\in l^2(R)$，则不存在 $\boldsymbol{f}\in l^2(R)$ 使得 $\boldsymbol{D}(\boldsymbol{f})=\boldsymbol{g}$，所以算子方程可能无解。如取 $\boldsymbol{f_n}\in l^2(\boldsymbol{R}), \boldsymbol{f_n}(j)=\delta_{n,j}$ 有

$$D(f_n)=\left(\cdots\frac{1}{j}\delta_{n,j}\cdots\right),\quad \|D(f_n)\|=\frac{1}{n}\to 0(n\to\infty)$$

则 f_n 并不趋近于 0，所以稳定性也不成立。

如果取 $D:H_1\to H_2, H_1=l^2(R), H_2=\left\{(f_1,f_2\cdots),\sum_{j=1}^{\infty}j^2f_j{}^2<\infty\right\}$，则这个算子是单射，同时也是满射。

设 $(f_1,f_2\cdots f_n\cdots)\in\mathrm{Range}(D)$，则有

$$(f_1,2f_2,\cdots if_i,\cdots)\in l^2(R),$$

$$D(f_1,2f_2,\cdots if_i\cdots)=(f_1,f_2,\cdots f_i,\cdots)$$

设

$$f=(f_1,f_2,\cdots f_i,\cdots)\in l^2(R),\ D(f)=\left(f_1,\frac{1}{2}f_2,\cdots,\frac{l}{i}f_i\cdots\right)\ \|Df\|_{H_2}=\sum_{i=1}^{\infty}i^2\cdot(\frac{l}{i}f_i)^2=\sum_{i=1}^{\infty}f_i{}^2=\|f\|_{H_1}$$

则 $\|D(f_1-f_2)\|_{H_2}=\|f_1-f_2\|_{H_1}$ 满足稳定性条件。

定义 1.1.2 有界线性算子 $K:H_1\to H_2$ 称为一个紧算子，当且仅当 H_1 中任一有界集合的像集合的闭包是 H_2 中的紧子集。也就是说对于 H_1 中的每个有界序列 $\{x_n\}$，由线性算子 K 作用得到的序列 $\{K(x_n)\}$ 有收敛子序列，则称 K 为紧算子。

例 1.1.2 设 $K:H_1\to H_2$ 是一个线性算子，如果 $\mathrm{range}(K)$ 是有限维空间，那么 K 是一个紧算子，特别地，矩阵作为线性变换（算子）也是紧算子。

例 1.1.3 $K(f)=\int_{\Omega}K(x,y)f(y)\mathrm{d}y$ 是 $L^2(\Omega)$ 上的紧算子。

下面的定理揭示了紧算子和算子方程病态性的关系。

定理 1.1.1 假设 $K:H_1\to H_2$ 是一个紧线性算子，H_1 和 H_2 都是无限维空间。如果 $\mathrm{Range}(K)$ 是无限维的，那么算子方程 $Kf=g$ 则不满足适定性定义的条件（1）和条件（3）（满射和稳定性），所以是一个病态方程。如果 $\mathrm{Range}(K)$ 是有限维的，那么算子方程 $Kf=g$ 则不满足适定性定义的条件（2）（单射），所以是一个病态方程。

1.2 最优化理论

定义 1.2.1 假设 $J:X\to R$ 是一个泛函，C 是 X 的子集，如果存在 $\delta>0$ 使得对于 $x\in C$，$\|x-x^*\|<\delta$ 时，有 $J(x^*)\leqslant J(x)$，则称 x^* 是 $J(x)$ 在 C 上的一个局部极小值点，记为 $x^*=\arg\min\limits_{x\in C}J(x)$。

定义 1.2.2 设 $J:X\to R$，如果对任意 $x,y\in X$ 和 $0<\tau<1$ 满足

$$J(\tau x+(1-\tau)y)\leqslant\tau J(x)+(1-\tau)J(y)$$

则称 $J(x)$ 是 X 上的凸函数。

定义 1.2.3 设 $\{x_n\}$ 是任意一个实数列，它的所有收敛子数列（容许收敛到 $\pm\infty$）的极限值

之最小数（最大数）称为数列$\{x_n\}$的下极限（上极限），记作$\liminf\limits_{n\to\infty} x_n$（$\limsup\limits_{n\to\infty} x_n$）。

定义 1.2.4 设$J:X\to R$是一个泛函，对于每个收敛于x^*的序列$\{x_n\}$，总有

$$J(x^*)\leqslant \liminf_{n\to\infty} J(x_n)$$

则称$J(x)$在x^*处是下半连续的，如果$J(x)$在定义域的每一点都是下半连续的，则称$J(x)$是下半连续的。

定理 1.2.1 对于函数$f:R^n\to[-\infty,+\infty]$以下各条件等价：

（1）水平集$V_r=\{x|f(x)\leqslant r\}$对每个$r$均为闭集；

（2）函数f为下半连续；

（3）集合$epif$为闭集。

证明：如果$f(x)=\infty$，那么V_r是空集，命题成立。先假设至少有一个$x\in R^n$，使$f(x)<\infty$，这样$epif$非空，且f至少有一个非空的水平集。

由条件（1）$\Rightarrow$条件（2）假设对于每个r，V_r为闭集，反设对于某个$\overline{x}$和收敛到$\overline{x}$的序列$\{x_k\}$有

$$f(\overline{x})>\liminf_{k\to\infty} f(x_k)$$

取γ满足$f(\overline{x})>\gamma>\liminf\limits_{k\to\infty} f(x_k)$，那么必存在子序列$\{x_{k_i}\}$，使得对所有$i\in I$，$f(x_{k_i})\leqslant\gamma$。这是因为如果对所有子序列，都有$f(x_{k_i})>\gamma$，那么$\liminf f(x_{k_i})\geqslant\gamma$，这与$\gamma$的取法相矛盾。由于$f(x_{k_i})\leqslant\gamma$，从而有$\{x_{k_i}\}\in V_\gamma$，又由于$V_\gamma$是闭集，$x_{k_i}\to\overline{x}$，于是$\overline{x}\in V_\gamma$，$f(\overline{x})\leqslant\gamma$。这与$\gamma$的选取矛盾。

由条件（2）$\Rightarrow$条件（3）假设f在R^n下半连续，并令$(\overline{x},\overline{\omega})$为点列$\{(x_k,\omega_k)\}\in epif$的极限点，由$epif$的定义知$f(x_k)\leqslant\omega_k$，由于$f$在$\overline{x}$下半连续，所以

$$f(\overline{x})\leqslant\liminf_{k\to\infty} f(x_k)\leqslant\overline{\omega}$$

即$(\overline{x},\overline{\omega})\in epif$，故$epif$为闭集。

由条件（3）$\Rightarrow$条件（1）假设$epif$为闭集，令$\{x_k\}$为点列，它收敛于某个$\overline{x}$，且属于某个水平集V_r，于是$(x_k,r)\in epif$，$(x_k,r)\to(\overline{x},r)$。由于$epif$为闭集，所以$(\overline{x},r)\in epif$，即$f(\overline{x})\leqslant r$，$\overline{x}\in V_r$，故$V_r$是闭集。证毕。

设$u=(u_1,u_2,\cdots,u_n)^{\mathrm{T}}$，其$l_1$范数$\|u\|_1=\sum\limits_{i=1}^{n}|u_i|$是下半连续的，因为集合

$$S=\left\{u\in R^n\,\middle|\,\|u\|_1\leqslant\lambda\right\}$$

是闭集，如果$\lambda<0$，那么S是空集，自然是闭集。如果$\lambda\geqslant 0$，设有序列$\{u_n\}$收敛于$u\in R^n$，且$u_n\in S$，那么

$$\|u\|_1=\|u-u_n+u_n\|_1\leqslant\|u-u_n\|_1+\|u_n\|_1\leqslant\|u-u_n\|_1+\lambda$$

两边关于n取极限，得

$$\|u\|_1 \leqslant \limsup_{n\to 0}(\|u-u_n\|_1+\lambda) \leqslant \lambda$$

即 $u \in S$，S 是闭集，从而该范数是下半连续的。

定义 1.2.5 泛函 $J: X \to R$，则称其是强制的，如果对于任意 $u_n \in X$，当 $\|u_n\| \to \infty$ 时，有 $J(u_n) \to \infty$。

在有限维空间，有界序列都有收敛的子序列，但在无限维空间，有界序列不一定有收敛的子序列，这就引出了弱收敛的概念。

定义 1.2.6 设 $\{x_n\}$ 是 Hilbert 空间 X 的一个序列，如果存在 x^* 使得对任意 $x \in X$，有 $\lim\limits_{n\to\infty}\left\langle x_n - x^*, x\right\rangle_X = 0$，则称 x_n 弱收敛于 x^*。

例 1.2.1 设 $f_n(x) = \dfrac{1}{\sqrt{2}}\sin(2n\pi x)$, $(n=1,2,\cdots)$，则 $\{f_n(x)\}$ 是 $L^2(0,1)$ 上的正交集。对任意 $f(x) \in L^2(0,1)$，有

$$\int_0^1 f_n(x)\cdot f(x)\mathrm{d}x = \frac{1}{\sqrt{2}}\int_0^1 \sin(2n\pi x)f(x)\mathrm{d}x$$

$$=\frac{1}{\sqrt{2}}\cdot\frac{1}{2\pi}\int_0^{2\pi}\sin(ny)f\left(\frac{y}{2\pi}\right)\mathrm{d}y\text{，（}y=2\pi x\text{）}$$

$\dfrac{1}{2\pi}\displaystyle\int_0^{2\pi}\sin(ny)f\left(\frac{y}{2\pi}\right)\mathrm{d}y$ 是 $f\left(\dfrac{y}{2\pi}\right)$ 的傅里叶级数的系数，由勒贝格引理知

$$\text{当 } n\to\infty \text{ 时，} \frac{1}{2\pi}\int_0^{2\pi}\sin(ny)f\left(\frac{y}{2\pi}\right)\mathrm{d}y \to 0$$

$$\text{即当 } n\to\infty \text{ 时，} \langle f_n(x)-0, f\rangle \to 0\text{，}$$

也就是 $f_n(x)$ 弱收敛于 0，但 $\|f_n(x)-0\|^2 = \|f_n(x)\|^2 = \dfrac{1}{2}\displaystyle\int_0^1(\sin(2n\pi x))^2\mathrm{d}x = 1$。

如果序列 $\{x_n\}$ 弱收敛于 x^*，且满足 $J(x^*) \leqslant \liminf\limits_{n\to\infty} J(x_n)$，则称 $J(x)$ 在 x^* 处是弱下半连续的。

在有限维空间，有界闭集就是紧集，但在无限维空间，这个关系却不一定成立。

定理 1.2.2 巴拿赫空间 X 是有限维的，当且仅当 X 中的单位闭球是紧集。

证明：令 $B=\{u\in X, \|u\|\leqslant 1\}$，如果 $\dim X < \infty$，那么有限维空间的有界闭集是紧集，因而 B 是紧集。

如果 $\dim X = \infty$，X 是一个可分的 Hilbert 空间，那么存在可数个单位正交系 $\{u_n\}\in X$，当 $m \neq n$ 时，有

$$\|u_n - u_m\|^2 = \|u_n\|^2 + \|u_m\|^2 = 2$$

序列 $\{u_n\}\in B$，但是它没有收敛的子序列，因而 B 不是紧集。

当 X 是一般的无限维巴拿赫空间时，设 W 是 X 中的闭线性子空间且 $W \neq X$，取 $v \in X - W$，那么 $\mathrm{dist}(v,W) > 0$，否则由下确界的定义可知，存在 W 中的序列 $\{w_n\}$ 使得当 $n\to\infty$ 时，$\|v-w_n\|\to 0$。由于 W 是闭集，所以 $v\in W$ 与 $v \in X-W$ 矛盾。

取 $w_\varepsilon \in W$，使得对任意的 $\varepsilon\in(0,1)$，$0<\|v-w_\varepsilon\|<(1-\varepsilon)^{-1}\mathrm{dist}(v,W)$，令 $u_\varepsilon = \dfrac{v-w_\varepsilon}{\|v-w_\varepsilon\|}$，则

$$\|u_\varepsilon\|=1\text{，且 }\operatorname{dist}(u_\varepsilon,W)\geqslant 1-\varepsilon$$

这是因为对所有的 $w\in W$，有

$$\|u_\varepsilon-w\|=\|v-w_\varepsilon\|^{-1}\left\|v-w_\varepsilon-\|v-w_\varepsilon\|w\right\|\geqslant\|v-w_\varepsilon\|^{-1}\operatorname{dist}(v,W)\geqslant 1-\varepsilon$$

取 $w_1\in X$，且 $\|w_1\|=1$，记 $W=\operatorname{span}\{w_1\}$，那么存在 $w_2\in X$，$\|w_2\|=1$ 且使 $\|w_1-w_2\|>2^{-1}$。再取 $W=\operatorname{span}\{w_1,w_2\}$，那么存在 $w_3\in X$，$\|w_3\|=1$ 且使 $\operatorname{dist}(w_3,W)\geqslant 2^{-1}$，即 $\|w_3-w_1\|>2^{-1}$，$\|w_3-w_2\|>2^{-1}$，把这种构造继续下去，则得 X 中的序列 $\{w_n\}$，$\|w_n\|=1$，当 $m\neq n$ 时，$\|w_n-w_m\|>2^{-1}$，因此，$\{w_n\}\in B$ 没有收敛的子序列，因而 B 不是紧集。

定理 1.2.3 K 上的 Hilbert 空间 X 中的每个有界序列 $\{u_n\}$ 都有一个弱收敛的子序列，K 是实数域或复数域。

证明：当 $X=\{0\}$ 时，定理结果显然成立。

（1）假设 X 是一可分的 Hilbert 空间，取 $\{v_n\}$ 为在 X 中稠密的可数集，对所有的 n，由于 $|(u_n|v_1)|\leqslant\|u_n\|\|v_1\|$，所以序列 $(u_n|v_1)$ 是 K 中的有界集，从而该序列有收敛的子序列，记为 $(u_{1n}|v_1)$，当 $n\to\infty$ 时，$(u_{1n}|v_1)\to a_1$。

对所有的 n，由于 $|(u_{1n}|v_2)|\leqslant\|u_{1n}\|\|v_2\|$，所以 $(u_{1n}|v_2)$ 是 K 中的有界集，从而该序列有收敛的子序列，记为 $(u_{2n}|v_2)$，当 $n\to\infty$ 时，$(u_{2n}|v_2)\to a_2$。以此类推，则有

$$\begin{matrix}(u_{11}|v_1), & (u_{12}|v_1), & (u_{13}|v_1) & \cdots & \to & a_1\\ (u_{21}|v_2) & (u_{22}|v_2) & (u_{23}|v_2) & \cdots & \to & a_2\\ \vdots & \vdots & \vdots & \cdots & & \vdots\end{matrix}$$

记 $w_n=u_{nn}$，那么对所有的 k，当 $n\to\infty$ 时，$(w_n|v_k)\to a_k$。
对每个 $v\in X$，存在序列 $\{v_k\}$，当 $k\to\infty$ 时，$v_k\to v$，那么

$$|(w_n-w_m|v)|=|(w_n-w_m|v-v_k+v_k)|\leqslant\|w_n-w_m\|\|v-v_k\|+|(w_n|v_k)-(w_m|v_k)|$$

其中，$\|w_n-w_m\|$ 是有界量。当 k 充分大时，$\|v-v_k\|$ 是无穷小，当 m,n 充分大时，$|(w_n|v_k)-(w_m|v_k)|$ 也是无穷小，所以 $\{(w_n|v)\}$ 是基本列，从而当 $n\to\infty$ 时，$(w_n|v)\to a(v)$。

定义 X 上的映射：$v\to a(v)$，显然它是一个线性有界泛函，由于 $a(v)=\lim\limits_{n\to\infty}(w_n|v)$，所以有

$$|a(v)|=\lim_{n\to\infty}|(w_n|v)|\leqslant\|v\|\overline{\lim_{n\to\infty}}\|w_n\|\leqslant\|v\|\sup_n\|w_n\|$$

由 Riesz 表示定理知，存在 $w\in X$，使得对所有 $v\in X$，有

$$a(v)=(w|v)$$

从而当 $n\to\infty$ 时，$(w_n|v)\to a(v)=(w|v)$，即 w_n 弱收敛于 w。

（2）如果 X 不是可分的 Hilbert 空间，记 Y 是由 $\{u_1,u_2,\cdots,u_n,\cdots\}$ 所张成子空间的闭包，即 $Y=\overline{\operatorname{span}\{u_1,u_2,\cdots u_n,\cdots\}}$，从而 Y 是可分的。由（1）可知，存在 $\{u_n\}$ 的子序列 $\{w_n\}$ 使得 w_n 弱收敛于 $w\in Y$，即对于所有的 $v\in Y$，当 $n\to\infty$ 时，$(w_n|v)\to(w|v)$。

对于 $\forall z \in X$， $z = v + v^{\perp}$， $v \in Y, v^{\perp} \in Y^{\perp}$， 有

$$(v^{\perp} | w) = 0 \ (v^{\perp} | w_n) = 0$$

进而有 $(z|w) = (v + v^{\perp}|w) = (v|w) + (v^{\perp}|w) = (v|w)$

$$(z|w_n) = (v + v^{\perp}|w_n) = (v|w_n) + (v^{\perp}|w_n) = (v|w_n)$$

当 $n \to \infty$ 时，$(w_n|v) \to (w|v)$，即 $(w_n|z) \to (w|z)$。

定理 1.2.4 假设泛函 $J: M \to R$ 具有下列性质：

（1）M 是实 Hilbert 空间 X 的非空闭凸子集；

（2）J 是弱下半连续的；

（3）如果集合 M 无界，那么 J 是强制的。

如果 $\min\limits_{u \in M} J(u)$ 有解，J 是严格凸泛函，那么极小值是唯一的。

证明：首先假设 M 是有界的，记 $\gamma = \inf\limits_{u \in M} J(u)$，$-\infty \leqslant \gamma < +\infty$，由下确界的定义可知，存在 M 中的序列 $\{u_n\}$，使得当 $n \to \infty$ 时，$J(u_n) \to \gamma$。由于 M 是有界的，从而序列 $\{u_n\}$ 也有界，由定理 1.2.3 可知 $\{u_n\}$ 有弱收敛的子序列 $\{u_{1n}\}$，使得 $\{u_{1n}\}$ 弱收敛于 u_*，由于 u_* 含于 $\{u_1, u_2, \cdots, u_n, \cdots\}$ 的闭凸包中（在 Hilbert 空间的闭凸集中，弱收敛序列的弱极限仍在此闭凸集中），从而有 $u_* \in M$，又因为 J 是弱下半连续的，所以有

$$J(u_*) \leqslant \lim_{n \to \infty} \inf J(u_n) = \gamma$$

即 $J(u_*) = \gamma$。

如果 M 是有无界集合，取 $v \in M$，由于 J 是强制的，当 $\|u\| \to \infty$ 时，$J(u) \to \infty$，于是存在 $r > 0$，使得对所有 $u \in M$，当 $\|u\| > r$ 时，$J(u) > J(v)$。

令 $M_r = \{u \in M, \|u\| \leqslant r\}$，$M_r$ 非空，$\min\limits_{u \in M_r} J(u)$ 的每个解都是 $\min\limits_{u \in M} J(u)$ 的解。由于 M_r 是闭凸的有界集，从而 $\min\limits_{u \in M_r} J(u)$ 有解，进而 $\min\limits_{u \in M} J(u)$ 有解。

如果 $J(u)$ 是严格凸函数，且 $J(u_0) = \gamma$， $u_* \neq u_0$，那么有

$$J(\frac{u_0}{2} + \frac{u_*}{2}) < \frac{1}{2} J(u_0) + \frac{1}{2} J(u_*) = \gamma$$

矛盾，从而有唯一的极小点。

定义 1.2.7 泛函 $J: X \to R$ 是凸的，$u \in X$，若存在 $p \in X^*$（X 的对偶空间）使得对任意的 $v \in X$，有

$$J(v) - J(u) - < p, v - u > \geqslant 0$$

则称 p 为 J 在 u 的次梯度。

J 在 u 的所有次梯度的全体称为 J 在 u 的次微分，记为 $\partial J(u)$。

例如 $\boldsymbol{f}(\boldsymbol{x}) = |x|, x \in \boldsymbol{R}$，在 $x = 0$ 处不可微，$\partial f(0) = [-1, 1]$。

再如 $\boldsymbol{f}(x) = \|\boldsymbol{x}\|_2, \boldsymbol{x} \in \boldsymbol{R}^n$，同样在 $\boldsymbol{x} = 0$ 处不可微，即

$$\partial f(\boldsymbol{x})=\begin{cases} B & , \boldsymbol{x}=0 \\ \dfrac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2} & , \boldsymbol{x}\neq 0 \end{cases}$$

其中，B 为以原点为圆心的单位闭球。

定义 1.2.8 设有泛函 $J:X\to R$，如果当 $\lambda\to 0^+$ 时

$$\lim_{\lambda\to 0^+}\frac{J(u+\lambda v)-J(u)}{\lambda}$$

存在，则称此极限为 J 在 u 点沿方向 v 的方向导数，记为 $J'(u;v)$。

如果存在 $p\in X^*$ 使得对任意 $v\in X$ 都有 $J'(u;v)=<p,v>$，则称 J 在 u 点是 Gateaux 可导的，p 称为 J 在 u 点 Gateaux 导数，记为 $J'(u)$。

定义 1.2.9 算子 $J:H_1\to H_2$ 在 u 处，$(u\in H_1)$ 是 Frechet 可导的，当且仅当存在 $J'(u)\in L(H_1, H_2)$，且满足

$$J(u+h)=J(u)+J'(u)h+o(\|h\|_{H_1})$$

或

$$\lim_{\|h\|\to 0}\frac{\|J(u+h)-J(u)-J'(u)h\|_{H_2}}{\|h\|_{H_1}}=0$$

则称 $J'(u)$ 为 J 在 u 的 Frechet 导数。

如果 $J:R^n\to\overline{R}$ 在 $\overline{\boldsymbol{x}}\in\operatorname{int}(\operatorname{dom}f)$ 处是 Frechet 可导的，其 Frechet 导数为 $\nabla f(\overline{\boldsymbol{x}})$，且有

$$\lim_{\|\boldsymbol{h}\|\to 0}\frac{|J(\boldsymbol{u}+\boldsymbol{h})-J(u)-<\nabla f(\overline{\boldsymbol{x}}),\boldsymbol{h}>|}{\|\boldsymbol{h}\|}=0$$

定理 1.2.5 设 $f:R^n\to\overline{R}$ 是任一函数，它在 $\overline{\boldsymbol{x}}\in\operatorname{dom}(f)$ 附近是局部 Lipschitz 连续的，则下列结论是等价的：

（1）f 在 $\overline{\boldsymbol{x}}$ 处是 Frechet 可微的；

（2）f 在 $\overline{\boldsymbol{x}}$ 处是 Gateaux 可微的。

证明：设 f 在 $\overline{\boldsymbol{x}}$ 处是 Frechet 可微的，对于取定的 $0\neq\boldsymbol{v}\in R^n$，则

$$\lim_{t\to 0}\frac{f(\overline{\boldsymbol{x}}+t\boldsymbol{v})-f(\overline{\boldsymbol{x}})-t<\nabla f(\overline{\boldsymbol{x}}),\boldsymbol{v}>}{t}=\|\boldsymbol{v}\|\lim_{t\to 0}\frac{f(\overline{\boldsymbol{x}}+t\boldsymbol{v})-f(\overline{\boldsymbol{x}})-<\nabla f(\overline{\boldsymbol{x}}),t\boldsymbol{v}>}{t\|\boldsymbol{v}\|}=0$$

这就蕴含着 f 在 $\overline{\boldsymbol{x}}$ 处是 Gateaux 可微的，且 $f'(\overline{\boldsymbol{x}})=\nabla f(\overline{\boldsymbol{x}})$，从这个推导还可以看出 f 在 $\overline{\boldsymbol{x}}$ 处 Frechet 可微蕴含着 f 在 $\overline{\boldsymbol{x}}$ 处 Gateaux 可微。

现在假设 f 在 $\overline{x}$ 处是 Gateaux 可微的，$f'(\overline{\boldsymbol{x}})=\boldsymbol{u}$，$|f(\overline{\boldsymbol{x}}+\boldsymbol{h})-f(\overline{\boldsymbol{x}})|\leqslant l\|\boldsymbol{h}\|$，下证

$$\lim_{\|\boldsymbol{h}\|\to 0}\frac{|f(\overline{\boldsymbol{x}}+\boldsymbol{h})-f(\overline{\boldsymbol{x}})-<\boldsymbol{u},\boldsymbol{h}>|}{\|\boldsymbol{h}\|}=0$$

反设上式不成立，则可找到 $\varepsilon_0>0$，序列 $\boldsymbol{h}_k\to 0$ 且 $\boldsymbol{h}_k\neq 0$，使得

$$\frac{|f(\overline{\boldsymbol{x}}+\boldsymbol{h}_k)-f(\overline{\boldsymbol{x}})-<\boldsymbol{u},\boldsymbol{h}_k>|}{\|\boldsymbol{h}_k\|}\geqslant\varepsilon_0$$

记 $t_k=\|\boldsymbol{h}_k\|$，$\boldsymbol{d}_k=\dfrac{\boldsymbol{h}_k}{\|h_k\|}$，则 $t_k\to 0$。不失一般性，可设 $\boldsymbol{d}_k\to\boldsymbol{d}$ 且 $\|\boldsymbol{d}\|=1$，因此有估计式

$$\varepsilon_0\leqslant\frac{|f(\bar{\boldsymbol{x}}+t_k\boldsymbol{d}_k)-f(\bar{\boldsymbol{x}})-<\boldsymbol{u},t_k\boldsymbol{d}_k>|}{t_k}$$

$$\leqslant\frac{|f(\bar{\boldsymbol{x}}+t_k\boldsymbol{d}_k)-f(\bar{\boldsymbol{x}}+t_k\boldsymbol{d})|+|<\boldsymbol{u},t_k\boldsymbol{d}>-<\boldsymbol{u},t_k\boldsymbol{d}_k>|}{t_k}$$

$$+\frac{|f(\bar{\boldsymbol{x}}+t_k\boldsymbol{d})-f(\bar{\boldsymbol{x}})-<\boldsymbol{u},t_k\boldsymbol{d}>|}{t_k}$$

$$\leqslant l\|\boldsymbol{d}_k-\boldsymbol{d}\|+\|\boldsymbol{u}\|\cdot\|\boldsymbol{d}_k-\boldsymbol{d}\|+\frac{|f(\bar{\boldsymbol{x}}+t_k\boldsymbol{d})-f(\bar{\boldsymbol{x}})-<\boldsymbol{u},t_k\boldsymbol{d}>|}{t_k}\to 0,\quad (k\to\infty)$$

这是一个矛盾，因此 f 在 $\bar{\boldsymbol{x}}$ 处是 Frechet 可微的。

定理 1.2.6 设 $f:R^n\to\bar{R}$ 是凸函数且 $\bar{\boldsymbol{x}}\in\mathrm{int}(\mathrm{dom}f)$，则下列断言是等价的：

（1）f 在 $\bar{\boldsymbol{x}}$ 处是 Frechet 可微的；

（2）f 在 $\bar{\boldsymbol{x}}$ 处是 Gateaux 可微的；

（3）$\partial f(\bar{\boldsymbol{x}})$ 是一单点集。

定义 1.2.10 假设 $J:H\to R$ 在 u 处 Frechet 可导的，$J'(u)\in L(H,R)$ 是一个泛函，由 Hilbert 空间的 Riesz 表示定理，H 空间上的任一泛函表示为内积的形式，即存在 $\mathrm{grad}J(u)\in H$，称为 $J(u)$ 的梯度，使得 $J'(u)h=\langle\mathrm{grad}J(u),h\rangle_H$。

定理 1.2.7 如果 $J:H\to R$ 在 u 处 Frechet 可导的，那么对任意 $h\in H$，映射 $\tau\to J(u+\tau h)$ $(R\to R)$ 在 $\tau=0$ 处可微，且

$$\frac{\mathrm{d}}{\mathrm{d}\tau}J(u+\tau h)\Big|_{\tau=0}=\langle\mathrm{grad}J(u),h\rangle_H\text{。}$$

例 1.2.2 设 $K:H_1\to H_2$ 是一个有界线性算子，固定 $g\in H_2$ 和 $\alpha>0$，求

$$J(u)=\frac{1}{2}(\|Ku-g\|_{H_2}^2+\alpha\|u\|_{H_1}^2)$$

梯度 $\mathrm{grad}J(u)$。

解：$\dfrac{\mathrm{d}}{\mathrm{d}\tau}J(u+\tau h)\Big|_{\tau=0}=\langle Ku-g,Kh\rangle_{H_2}+\alpha\langle u,h\rangle_{H_1}$

$$=\langle K^*(Ku-g),h\rangle_{H_1}+\langle\alpha u,h\rangle_{H_1}$$

$$=\langle(K^*K+\alpha I)u-K^*g,h\rangle_{H_1}$$

从而 $\mathrm{grad}J(u)=(K^*K+\alpha I)u-K^*g$。

定理 1.2.8 设 $J:H\to R$ 是 Frechet 可微的，如果 $J(u)$ 在 u^* 处取得局部无约束极小值，那么 $\mathrm{grad}J(u^*)=0$。

证明：取 $g=\mathrm{grad}J(u^*)$，$J(u^*-\tau g)=J(u^*)-\tau J'(u^*)g+o(\tau\|g\|)$，有

$$J'(u^*)g=\langle\mathrm{grad}J(u^*),g\rangle=\|g\|^2,\quad J(u^*-\tau g)=J(u^*)-\tau\|g\|^2+o(\tau)$$

取 $\tau>0$ 充分小，若 $g\neq 0$，$J(u^*-\tau g)<J(u^*)$，矛盾，从而 $g=0$。

$J:R^n\to R$ 具有连续的二阶偏导数，定义其二阶 Frechet 导数为

$$\mathbf{Hess}J(\boldsymbol{u})=\left(\frac{\partial^2 J}{\partial u_i \partial u_j}\right)_{n\times n}，\ 1\leqslant i,j\leqslant n$$

定理 1.2.9 设 J 在凸集 C 上二次 Frechet 可微，J 是凸函数当且仅当 $\mathrm{Hess}J(u)$ 在 C 内部是半正定的，如果 $\mathrm{Hess}J(u)$ 是正定的，则 $J(f)$ 是严格凸的。

假设 $J(u)$ 是二次 Frechet 可微的，那么有泰勒展开式

$$J(u+h)=J(u)+\langle \mathrm{grad}J(u),h\rangle+\frac{1}{2}\langle \mathrm{Hess}J(u)h,h\rangle+o(\|h\|^2)$$

定理 1.2.10 如果 $J(u)$ 在 u_* 处有二阶 Frechet 导数，且 $\mathrm{grad}J(u_*)=0$，$\mathrm{Hess}J(u_*)$ 正定，那么 u_* 是 $J(u)$ 的局部极小点。

1.3 泛函的变分、Euler–Lagrange 方程和边界条件

基于变分法的图像复原是一种确定性的方法，它通过引入能量函数，将图像复原问题转化成泛函极值问题，再由变分法的相关理论将其转化为求解一个线性方程组或偏微分方程。简单地说，它可以分为以下几个主要步骤：

（1）从实际问题的物理意义出发，建立相应的泛函极小值问题并寻找其约束条件；

（2）利用变分法求得极小问题的 Euler-Lagrange 方程，该方程通常是一个线性方程组或偏微分方程；

（3）在适当的边界条件下求 Euler-Lagrange 方程的数值解。

设有如下的泛函极小问题的

$$\min J(u(x,y))=\min\iint_{\Omega}F(x,y,u,p,q,r,s,t)\mathrm{d}x\mathrm{d}y$$

其中

$$p=\frac{\partial u}{\partial x},q=\frac{\partial u}{\partial y}，\quad r=\frac{\partial^2 u}{\partial x^2},s=\frac{\partial^2 u}{\partial x\partial y},t=\frac{\partial^2 u}{\partial y^2}$$

该泛函的一阶变分为

$$\begin{aligned}\delta J(u(x,y))&=\left.\frac{\mathrm{d}J(u+\alpha\delta u)}{\mathrm{d}\alpha}\right|_{\alpha=0}\\&=\iint_{\Omega}(F_u-(F_p)_x-(F_q)_y)\delta u\mathrm{d}x\mathrm{d}y+\iint_{\Omega}((F_p\delta u)_x+(F_q\delta u)_y)\mathrm{d}x\mathrm{d}y+\\&\quad\iint_{\Omega}(F_r(\delta u)_{xx}+F_t(\delta u)_{yy})\mathrm{d}x\mathrm{d}y+\iint_{\Omega}F_s(\delta u)_{xy}\mathrm{d}x\mathrm{d}y\end{aligned}$$

设 $\boldsymbol{n}=(n_1,n_2)$ 为边界 $\partial\Omega$ 上的外法线向量，利用格林公式