doi: 10.11720/wtyht.2021.0292

艾寒冰,王彦国. 基于 DCT 的广义指数阈值衰减凸集投影算法在位场数据补空中的应用[J].物探与化探,2021,45(6):1559-1568.http://doi. org/10.11720/wtyht.2021.0292

Ai H B, Wang Y G. Interpolation of potential-field data by Projection Onto Convex Sets algorithm with generalized exponential threshold and based on Discrete Cosine Transform [J]. Geophysical and Geochemical Exploration, 2021, 45(6):1559–1568.http://doi.org/10.11720/wtyht.2021.0292

# 基于 DCT 的广义指数阈值衰减凸集投影算法 在位场数据补空中的应用

# 艾寒冰,王彦国

(东华理工大学地球物理与测控技术学院,江西南昌 330013)

摘要:数据补空或插值是位场数据处理中一个非常基本且重要的问题,在野外数据测量过程中遇到水域、断崖等 情况时无法进行数据采集,导致数据出现缺失情况。为了便于后续数据处理,需要对缺失数据体进行补全。本文 将离散余弦变换(DCT)应用到凸集投影(POCS)算法中来补全位场缺失数据,并给出了广义指数阈值衰减方式。 模型试验及实例应用表明,基于 DCT 并结合广义指数阈值衰减形式的 POCS 算法具有数据补空精度高、补空痕迹 小和补空数据噪声含量更接近真实情况等优点。

# 0 引言

在野外重磁数据采集过程中,常会遇到沼泽、河流、水库或断崖等恶劣环境,导致数据无法采集,使得一部分数据空缺,不利于后期数据处理与解释。目前,解决这一问题的主要措施是进行数据网格插值,较为常用的有克里金插值法<sup>[1]</sup>、最小曲率法<sup>[2]</sup>、等效源法<sup>[3]</sup>、线性插值法<sup>[4]</sup>等,不过这些方法的数据补空精度较低<sup>[5]</sup>。王万银等<sup>[6]</sup>引入最小曲率插值方法,给出了其差分迭代格式,并讨论了相关技术措施,验证了该方法精度明显高于常规余弦插值法,但方法受给定初值影响较大;王明等<sup>[7]</sup>把位场数据插值类比成热流传导的过程,进而提出了基于热传导模型的位场网格数据补空方法,并验证该法比最小曲率法插值效果更优;闫浩飞等<sup>[5]</sup>采用基于线性阈值衰减的凸集投影算法(POCS)进行位场数据补

空,该方法相对常规插值方法具有更高的数据补空 精度;曾小牛等进一步将该方法进行推广,提出了基 于凸集投影算法的重力数据扩充、下延一体化方 法<sup>[8]</sup>和重力同时填充扩边和去噪方法<sup>[9]</sup>,有效地改 善了常规分步骤进行插值、扩边、下延和去噪的问 题,获得了良好的处理效果。

本文在凸集投影算法数据补空精度高基础上, 结合离散余弦变换(DCT)能量聚集度高及消耗内存 小的优点,并在 Gao 等<sup>[10]</sup>、张华等<sup>[11]</sup>在地震数据补 空的研究基础上,提出了基于 DCT 的广义指数阈值 衰减 POCS 算法,用于重磁数据的高质量补空工作。

### 1 理论基础

凸集投影算法广泛用于图像修复,该方法将所 有图像约束表示为希尔伯特空间中的一系列闭合凸 集 *C<sub>i</sub>*(*i*=1,2,3,…,*m*),然后将任意初始值 *x* 迭代投

#### 收稿日期: 2021-01-04; 修回日期: 2021-05-19

基金项目:国家重点研发计划项目"铀矿基地深部成矿条件地球物理探测技术研究"(2017YFC0602603)、江西省自然科学基金项目"基于位场 多源型 tilt-depth 法的场源深度快速自动反演研究"(20171BAB213030)、国家自然科学基金项目"基于位场广义梯度张量的欧拉反 褶积方法研究"(41504098)

通讯作者:王彦国(1985-),男,博士,副教授,硕士研究生导师,从事勘探重磁学方面的教学与研究工作。Email:wangyg8503@126.com

第一作者:艾寒冰(1998-),男,在读硕士研究生,主要从事重磁勘探方面的学习与研究工作。Email:1724178612@qq.com

影到投影算子 H 下的所有闭合凸集的交点上。将 位于交点边界上的最终投影点作为最优解,其原理 如图 1 所示, POCS 迭代过程表达式为:

$$x_{k+1} = H_m H_{m-1} \cdots H_1 x_k, \ k = 0, 1, \cdots, K_0$$
(1)



图 1 POCS 算法原理<sup>[10]</sup> Fig.1 The mechanism of POCS method<sup>[10]</sup>

基于 DCT 的数据补空凸集投影理论公式:  
$$D_k = D_{obs} + MC^{-1}T_kCD_{k-1}, k = 1,2,3,...,K,$$
 (2)

式中:K 是总迭代次数; $D_{obs}(x,y)$ 是含缺失数据的原 始观测数据; $D_k$  代表第 k 次迭代补空的数据;C 和  $C^{-1}$ 代表二维离散余弦正、反变换;M 是采样矩阵,其 与输入数据维度相同,如果某点值为0,代表该点有 数据,无需插值,反之为1,则代表该点需要进行插 值计算; $T_k$  是门限值矩阵,可表示为:

$$T_{k}(i,j) = \begin{cases} 1, & |S_{k}(i,j)| \ge p_{k} \\ 0, & |S_{k}(i,j)| < p_{k} \end{cases}$$
(3)

式中: $p_k$ 是第 k 次迭代阈值, $S_k$ 是第 k 次的 DCT 变换谱。

Abma<sup>[13]</sup>给出的线性阈值衰减形式为:

$$p_{k} = p_{k}^{\max} - \frac{(k-1)(p_{k}^{\max} - p_{k}^{\min})}{K-1}, \qquad (4)$$

式中: $p_k^{\max} = \max(\operatorname{abs}(S_k))$ ,  $p_k^{\min} = \min(\operatorname{abs}(S_k))$ ,  $k = 1, 2, \dots, K_{\circ}$ 

本文给出的广义指数阈值衰减形式为:

$$p_{k} = p_{k}^{\max} * \exp\left[-\left(\frac{k-1}{K-1}\right)^{\text{Para}} \ln\left(p_{k}^{\max}/p_{k}^{\min}\right)\right] (5)$$

式中:Para 为控制衰减速度的参数,当 Para = 1 时为 Gao 等<sup>[10]</sup>提出的阈值衰减情况,当 Para = 0.5 时即 为张华等<sup>[11]</sup>提出的阈值衰减形式。

凸集投影算法进行数据补空的基本流程为:对 原始输入数据采用 DCT 处理,给定一个阈值,使得 变换谱中振幅大于等于这个阈值时保留,振幅低于 这个阈值则充零代替,然后对该数据进行反离散余 弦变换乘以采样矩阵,最后将其与原始待补空数据 叠加,而在下一次迭代计算时,减小阈值重复上述操 作,直到达到结束迭代计算的条件。本算法需要实现对不规则数据网格化,以满足 DCT 变换的需要, 算法的关键在于选择合适的稀疏变换和阈值衰减方式,其工作流程图见图 2。



图 2 基于 DCT 的 POCS 数据补空流程 Fig.2 The flow-process diagram of POCS method to interpolate based on DCT

#### 2 模型试验

为了验证指数阈值衰减凸集投影算法在位场中 的应用效果及优越性,建立了一个由参数不同的4 个模型体产生的叠加重力异常进行试算,使用的模 型体参数如表1所示,模拟测区面积为500m×500 m,点距与线距均为10m。图3是完整的叠加重力 异常,其中黄色虚线所圈部分为地质体边缘在地面 的正投影,粉红色锯齿圈出区域为待补空部位。

图 4 展示了 Surfer 软件中 4 种常用的插值算法 所获得的数据补空结果及其与图 3 的残差情况。可 以看出,所有的常规插值算法都在模型体 1 处存在 等值线的间断,尤其反距离加权平均法的最为明显, 最小曲率法的等值线畸变最不明显,但最小曲率法 的残差值最大,其次是反距离加权法的。另外,常规 插值方法的残差均在原始数据的正异常处为负值, 主 1

古ヨシ新

Table 1 Model parameters						
Model 1	圆柱体	-200~200	-190~-110	10~90	2	
Model 2	棱柱体	-50~50	-50~50	20~400	-2	
Model 3	棱柱体	$-200 \sim 0$	100~200	30~1000	2	
Model 4	球体	100~200	100~200	10~110	2	



图 3 组合模型产生的重力异常及缺失部位(粉色部分)

Fig.3 Forward gravity anomaly with complete

data and the target zone (pink part) 而负异常处为正值,这表明了补空的数据没有达到 真实值的幅度。表2给出了4种方法的插值均方根 误差(RMS),可以看出,反距离加权法的数据补空 精度最差,其次是最小曲率法,二者误差是克里金法 和径向基函数法的2倍左右。

表 2 不同常规插值方法的数据补空误差

Table 2 RMS errors of interpolating data using different conventional interpolation methods

处理方法	均方根误差/mGal
克里金法	0.123
径向基函数法	0.113
反距离加权法	0.285
最小曲率法	0.217

为了了解线性阈值及指数阈值衰减凸集投影与 总迭代次数关系,图5绘制了不同模式下的凸集投 影方法数据补空误差与最大迭代次数K关系曲线。 可以看出,无论是线性衰减,还是不同参数的指数衰 减,数据补空误差均是随着总迭代次数 K 增加而逐 渐减小,不过指数衰减的速率要明显大于线性衰减, 其中 Para=0.5 时衰减速率最快。线性阈值衰减方 式 POCS 算法基本收敛在 0.070 mGal, 指数阈值衰减 基本收敛在 0.030 mGal, 不过指数阈值衰减的误差 随着迭代次数变化呈现出了一定的波动性。对比图 5 与表2 可以看出,线性阈值衰减、指数阈值衰减凸 集投影算法的数据补空精度分别是常规方法效果最 好的径向基函数法的1.6倍和4.5倍,这表明了采用 基于指数阈值衰减的凸集投影算法能够获得更好的 数据补空效果。

图 6 则是基于线性与指数阈值衰减的 POCS 算 法迭代 800 次数据补空后的重力异常及其与图 3 的 残差,可以看出,无论线性阈值,还是不同参数指数 阈值 POCS 法处理后的重力异常等值线均较圆滑, 并不存在明显的等值线畸变情况。从残差图中可以 看出.POCS 算法的残差在原始数据表现为正异常 区域内为正值,而负异常区为负值,这恰于常规插值 方法数据补空的残差是相反的,也就是说,基于 POCS 算法的数据补空的幅值略大于实际数据的。

为了进一步证实本文算法的有效性,对理论重 力异常(图3)添加了30%随机噪声(图7),然后再 对含噪重力异常进行了数据缺失处理(图7粉色部 位)。图8是4种常用插值方法获得的含噪重力数 据补空结果及其与无噪声重力异常的残差,表3给 出了4种插值方法的误差。对比表3与表2可以看 出,相对无噪声时,含噪数据的常规插值方法补空误 差有所增加,不过规律基本保持,还是反距离加权法 的误差最大,最小曲率法次之。从图 8 可以看出,4 种常规插值方法获得的补空数据仍在在模型体1位 置存在等值线的畸变情况,另外还可以看出插值区 域的噪声水平与周边数据噪声水平明显不一致,补 空区的等值线更加圆滑些。

图 9 是线性及指数阈值衰减 POCS 算法的补空 表 3 不同常规插值方法对含噪重力异常的数据补空误差 Table 3 RMS errors of interpolating noise-corrupted data using different conventional interpolation methods

插值方法	均方根误差/mGal
克里金法	0.191
径向基函数法	0.185
反距离加权法	0.315
最小曲率法	0.280

H



a, b—REE a the light of Kriging and its residuo; c, d—interpolated result of radial basis function and its residuo; e, f—interpolated result of inverse distance to a power and its residuo; g, h—interpolated result of minimum curvature and its residuo











a、b—线性阈值插值及残差; c、d—指数阈值(Para=0.5)及残差; e、f—指数阈值(Para=1)及残差; g、h—指数阈值(Para=2)及残差 a、b—interpolated result using linear threshold and its residuo; c、d—interpolated result using exponential threshold (Para=0.5) and its residuo; e、f—interpolated result using exponential threshold (Para=1) and its residuo; g、h—interpolated result using exponential threshold (Para=2) and its residuo 图 6 基于线性、指数阈值衰减的凸集投影算法数据补空结果及与原数据的残差(K=800)





图 7 图 3 添加 30%随机噪声后的重力异常及缺失部位 Fig.7 Gravity anomaly and missing site after adding 30% random noise of Fig. 3

的关系曲线。可以看出,线性阈值衰减 POCS 法的 补空数据仍然是随着总迭代次数的增加而减小,基 本收敛在 0.137 mGal,但不同 para 时的指数阈值衰 减 POCS 法数据插值误差则是随着迭代增加,先减 小后增加,最终基本收敛在 0.272 mGal 处,推测是误 差增加是由于噪声的反复重建导致的。图 10 可以 看出,线性阈值衰减的 POCS 法(图 10a)也在模型 体1处产生了明显的等值线间断情况,补空区内的 异常同样较为光滑,与残差图(图 10b)仍是南、北两 侧为大面积的正差值,而中部则整体表现为大面积 数据与无噪声理论值之间均方误差与总迭代次数 K 的负差值,这表明含噪数据进行补空时,线性阈值衰



a、b—克里金法插值及残差;c、d—径向基函数法插值及残差;e、f—反距离加权法插值及残差;g、h—最小曲率法及残差 a、b—interpolated result of Kriging and its residuo;c、d—interpolated result of radial basis function and its residuo;e、f—interpolated result of inverse distance to a power and its residuo;g、h—interpolated result of minimum curvature and its residuo

图 8 常规插值方法对含噪缺失数据体的补空结果及与原数据的残差



and the difference with the original noise-free data





Fig.9 The relationship between the errors of interpolation data using POCS with different threshold models for the noise-corrupted data and the total number of iterations



a、b—线性阈值插值及残差; c、d—指数阈值(Para=0.5)及残差; e、f—指数阈值(Para=1)及残差;g、h—指数阈值(Para=2)及残差 a、b—interpolated result using linear threshold and its residuo;c、d—interpolated result using exponential threshold (Para=0.5) and its residuo;e、f interpolated result using exponential threshold (Para=1) and its residuo;g、h—interpolated result using exponential threshold (Para=2) and its residuo

#### 图 10 图 7 采用线性与指数阈值衰减 POCS 算法的数据补空结果及其残差结果

## Fig.10 Results of interpolation using POCS method with linear and exponential threshold model and the differences with the noise-free data

减的 POCS 法数据填充的误差主要是数据未真实恢 复引起的。指数阈值衰减的 POCS 法(图 10c、e、g) 补空的数据明显也存在着局部的波动,即含有一定 的噪声,残差图(图 10d、f、h)则是正负异常间隔出 现,即认为主要误差是由噪声干扰引起的。

## 3 应用实例

为了检验本文算法的实际资料处理能力,对黑 龙江嫩北农场布格重力异常进行挖空与数据补空处 理。图 11 是网格密度 500 m×250 m 的布格重力异 常,布格重力异常总误差为 0.018 mGal。由于该区 中部 NWW 流向的科洛河在汛期时河床较宽,最宽 处可达 1 km,汛期时重力数据将无法采集,故这里 对科洛河的大致流域区域进行了数据挖空处理(图 11 粉色部位)。

图 12 是理论模型测试所提及的 4 种常用插值 方法处理的结果及其残差分布,可以看到,这 4 种插 值方法数据补空结果基本一致,很难看出彼此差异, 且等值线畸变现象也难以直接显示,其主要原因是 挖空区域宽度仅为 2~3 个点距。从残差图上可以 看出,补空数据与实际数据的残差仍是大面积的正、 负值,即误差具有一定的范围,这显然不利于后期数 据处理与解释。表4是的4种常规插值方法的数据 补空误差,可以看出,径向基函数法的数据补空误差 最小,其次是最小曲率法和克里金法,反距离加权法 的误差仍是最大,不过所有常规方法的数据补空误 差均远大于布格重力异常的总误差,也就是说补空 数据质量较差。

图 13 给出了基于线性、不同参数 Para 指数阈 值衰减 POCS 算法数据补空均方根误差与总迭代次 数 K 的关系曲线,可以看出,无论是线性、还是不同 参数的指数阈值衰减,POCS 算法数据重建误差均 随着 K 增加而逐渐减小,线性阈值衰减 POCS 算法 的误差基本收敛于0.149 mGal,远高于布格重力异



a、b—克里金法插值及残差;c、d—径向基函数法插值及残差;e、f—反距离加权法插值及残差;g、h—最小曲率法及残差

a b-interpolated result of Kriging and its residuo; c d-interpolated result of radial basis function and its residuo; e f-interpolated result of inverse distance to a power and its residuo; g h-interpolated result of minimum curvature and its residuo

#### 图 12 实际数据的常规插值方法数据补空结果

#### Fig.12 Results of interpolation using conventional methods and the differences with the real data





表 4 不同常规方法对实际资料进行数据补空的误差 Table 4 The errors of interpolation results using conventional methods for the read data

6期

处理方法名称	均方根误差/mGal
克里金法	0.105
径向基函数法	0.091
反距离加权法	0.298
最小曲率法	0.105

常总误差,即线性阈值衰减的 POCS 算法重建的数 据质量也较低;指数阈值衰减 POCS 算法的误差收 敛于 0.008 mGal,则低于布格重力异常总误差,重建 数据质量较高,重建数据完全可以替代真实数据。 图 14 给出了 Para=0.5 时指数阈值衰减 POCS 法的 处理结果以及残差分布,残差图中显示了缺失数据 宽度较小的西北侧基本为0,而补空区较宽的东南 侧基本是正负异常间隔出现,且数据幅值较小,这体 现出了本文方法在数据补空方面的优势。

# 4 结论

针对缺失的重磁数据需要有效恢复的问题,本 文采用了基于 DCT 的广义指数阈值衰减凸集投影 算法进行数据补空工作。通过模型分析与实例应 用,证实了基于指数阈值衰减的凸集投影算法在数 据补空中存在补空精度高及补空痕迹小的优点,另 外补空数据还能与实际数据噪声含量保持一致。



图 13 不同衰减方式 POCS 补空的误差与总迭代次数 K 的关系曲线

Fig.13 The relationship between the errors of interpolating data using POCS with different threshold models and the total number of iterations



#### 参考文献(References):

社,1990.

Sun H Q. Geological statistics and its application [M]. Beijing: China University of Mining and Technology Press, 1990.

[1] 孙洪泉.地质统计学及其应用[M].北京:中国矿业大学出版

- [2] Briggs I C. Machine contouring using minimum curvature [J]. Geophysics, 1974, 39(1): 39-48.
- [3] Cordell L. A scattered equivalent-source method for interpolation and gridding of potential-field data in three dimensions [J]. Geophysics, 1992, 57(4): 629-636.
- [4] 郭志宏.一种使用的等值线型数据网格化方法[J].物探与化探,2001,25(3):203-208.
  Guo Z H. A practical contour type data gridding technique [J].
  Geophysical and Geochemical Exploration, 2001, 25(3): 203-208
- [5] 闫浩飞,刘国峰,薛典军,等.基于凸集投影方法的重磁数据规则缺失重建[J].地球物理学进展,2016,31(5):2192-2197.
  Yan H F, Liu G F, Xue D J, et al. Reconstruction of gravity/magnetic data with the projection-onto-convex-sets methods [J]. Progress in Geophysics, 2016, 31(5): 2192-2197.
- [6] 王万银,邱之云,刘金兰,等.位场数据处理中的最小曲率扩边 和补空方法研究[J].地球物理学进展,2009,24(4):1327-1338.

Wang W Y, Qiu Z Y, Liu J L, et al. The research to the extending edge and interpolation based on the minimum curvature method in potential field data processing [J]. Progress in Geophysics, 2009, 24(4): 1327 – 1338.

[7] 王明,刘前坤,李芳,等.基于热传导模型位场网格数据补空方 法研究[J].物探与化探,2015,39(S1):144-151. Wang M, Liu Q K, Li F, et al. Filling grid dummy values by heat conduction model [J]. Geophysical and Geochemical Exploration, 2015, 39(S1): 144-151.

[8] 曾小牛,李夕海,刘继吴,等.基于凸集投影的重力数据扩充下 延一体化方法[J].石油地球物理勘探,2019,54(5):1166-1173.

Zeng X N, Li X H, Liu J H, et al. An integration of interpolation, edge padding, and downward continuation for gravity data based on the projection onto convex sets [J]. Oil Geophysical Prospecting, 2019, 54(4): 1166 – 1173.

- [9] 曾小牛,李夕海,侯维君,等.基于凸集投影的重力同时填充扩 边和去噪方法[J].石油地球物理勘探,2020,55(1):197-205. Zeng X N, Li X H, Hou W J, et al. Simultaneous interpolation, edge padding and denoising method for gravity data based on the projection onto convex sets [J]. Oil Geophysical Prospecting, 2020, 55(1): 197-205.
- [10] Gao J J, Chen X H, Li J Y, et al. Irregular seismic data reconstruction based on exponential threshold model of POCS method [J]. Applied Geophysics, 2010, 7(3): 229-238.
- [11] 张华,陈小宏.基于 jitter 采样和曲波变换的三维地震数据重建
  [J].地球物理学报,2013,56(5):1637-1649.
  Zhang H, Chen X H. Seismic data reconstruction based on jittered sampling and curvelet transform[J]. Chinese J. Geophys., 2013, 56(5):1637-1649.

# Interpolation of potential-field data by Projection Onto Convex Sets algorithm with generalized exponential threshold and based on Discrete Cosine Transform

#### AI Han-Bing, WANG Yan-Guo

(School of Geophysical and Measurement-Control Technology, East China University of Technology, Nanchang 330013, China)

**Abstract**: Data filling or interpolation is a fundamental and vital problem of potential-field data processing. Some data cannot be measured when some places are unable to reach, such as rivers, or cliffs. If we want to acquire the missing data for better subsequent processing, we need to interpolate or fill in the missing data. Hence, this article introduces the Discrete Cosine Transform (DCT) method into the Projection Onto Convex Sets (POCS) algorithm to tackle this problem, and a generalized exponential threshold attenuation method is also given. Finally, model tests and practical applications show that the POCS algorithm with generalized exponential threshold attenuation are closer to real situation.

Key words: data filling or interpolation; DCT; POCS; generalized threshold model

(本文编辑:王萌)