第 4 章　磁参数变化的 MT 二维反演问题

地球物理的反演方法是地球物理学科的一个重要组成部分，它对于解决地球物理中的由各种实测数据反映地球物理模型空间的问题起着至关重要的作用。反演的问题是建立在正演问题被解决之后的基础上的，著名的反演理论学者罗伯特·珀克（R.Parker，1970）曾把反演问题的研究归纳为四个方面的问题：解的存在性、模型构制、解的非唯一性和解的评价。这四个方面的问题就构成了地球物理反演的基本理论。在大地电磁测深中，对于 MT 资料的反演解释就是利用实际所测量的数据资料定量的求出其对应的地电断面的电性参数。具体说来，就是根据在地表观测的视电阻率、相位、阻抗或导纳，求从地壳表层到核幔边界的地下电性结构，确定出表征这些结构的特性参数。实际工作中比较普遍的是采用曲线拟合的方法，即寻找一个模型的理论曲线，让它与实测资料曲线进行对比，在经过多次的选择模型曲线进行对比之后两条曲线能够在均方差极小意义下最好的重合，这时的地电模型可以作为实测资料的反演解释结果。

目前主流的大地电磁反演中，我们仍是以电阻率作为单一的参数进行实测数据处理，对于多数的情况下这种方法是可行的。但在遇到一些高磁化率物质存在的地区时，仅仅靠反演电阻率参数而把磁参数当作常数进行数据处理，不只会浪费实测数据中的磁化率分布信息，而且也会导致资料解释的不够准确。我们可以从静态磁场数据中提取磁化率分布的信息，然后将这种信息作为一种先验信息加入到电磁数据反演中来获得电阻率的分布；或者是把通过其他手段获得的电阻率分布作为先验信息，反演电磁数据来得到磁化率的分布[37]。

本章将介绍一些 MT 中常用的二维反演方法，并对带磁化率参数和电阻率参数同时反演的计算和实现进行探讨。

4.1 大地电磁二维反演方法

大地电磁测深资料反演是大地电磁测深工作中最关键也是最困难的一环。因为大地电磁测深工作最终的成果集中体现在反演结果上，一切地质结论都是从反演结果上推论出来的，而大地电磁测深反演本身又存在着强烈的多解性，任何疏忽都会导致反演结果的错误甚至失败，使反演结果与客观实际大相径庭（刘国栋）。

虽然一维地质构造在实际工作中是很少遇到的，但由于一维反演的简单、快捷、比较成熟、正确处理下的结果比较接近客观实际等特点，目前仍在 MT 测深工作中广泛应用，而二维、三维等多维反演方法复杂，计算量大，有些方法仍处于研究和改进当中，还存在很多的问题与困难。因为在反演过程中，每迭代一次，对每个频率都要计算一次模型数据及其对该模型每个参数的偏导数，即要进行若

干次正演计算，对于多维计算来说，要用解析法求解反映各种地球物理场的偏微分方程非常困难。

假设定义模型空间中的一个点为 m，数据空间中的一个点为 d，那么反演问题可用下面的公式表示为

$$d = F(m) + e \tag{4－1}$$

式中 d 是数据向量，F 是正演模拟函数，m 是模型向量，e 是残差向量。其中 $d = \left[d^1 \quad d^2 \quad \cdots \quad d^N \right]^T$，分量 d^i 为 TE 或 TM 极化模式下复数视电阻率 ρ_{app} 取对数后的振幅或相位；$m = \left[m^1 \quad m^2 \quad \cdots \quad m^M \right]^T$ 为定义函数的参数向量，本文中因为考虑到磁化率参数，则 $m = \left[m_\rho^M , m_\kappa^M \right]^T$。

按照 Tikhonov 和 Arsenin(1977)的思路解反演问题，取“正则化解”为对给定的 λ、V 和 L 最小化如下目标函数 Ψ 的模型

$$\psi(m) = \left(d - F(m)\right)^T V^{-1} \left(d - F(m)\right) + \lambda m^T L^T L m \tag{4－2}$$

在大地电磁二维的反演方法中，常用的算法有 OCCAM 反演、快速松弛反演（RRI）、共轭梯度法、高斯－牛顿法、Mackie－Madden 法和非线性共轭梯度法等，下面介绍一些电阻率反演的方法，同时对磁化率为参数时进行一些相应公式的推导。

4.2 OCCAM 反演

OCCAM 反演（S.C.Constable 等，1987）实际是求一个多层地球模型的最光滑解，为了抑制不为数据所需要的模型结构，应最小化模型粗糙度。在一个二维构造中，设 y 轴为走向方向，则一阶导数和二阶导数模型粗糙度可分别表示为

$$R_1 = \left\| \partial_x m \right\|^2 + \left\| \partial_z m \right\|^2 \tag{4－3}$$

$$R_2 = \left\| \partial_x^2 m \right\|^2 + \left\| \partial_z^2 m \right\|^2 \tag{4－4}$$

式中 m 是模型电性参数，在本文中 m 分别为电阻率和磁化率的对数；∂_x、∂_z 各是一个粗糙化矩阵，分别为对侧向相邻单元的模型参数取差分和对垂向相邻单元的模型参数取差分。

由于模型网格在侧边为均匀层，在底部为均匀块，一阶导数光滑最好地符合了正演计算所要求的边界条件。假定二维有限元网格总共由 N 个单元构成，在垂直方向有 P 个单元，每个单元高度分别为 h_i，i＝1，2，…，P。在水平方向

有 H 个单元，每个单元的宽度为 w_i，$=1$，2，…，H；从左上角单元开始由左到右给单元编号，$N\times N$ 垂向粗糙化矩阵 ∂_z 可表示为

$$\partial_z=\begin{bmatrix} -1 & 0 & \cdots & 0 & 1 & 0 & 0 & \cdots \\ 0 & -1 & 0 & \cdots & 0 & 1 & 0 & \cdots \\ & & \ddots & & & & \ddots & \\ 0 & 0 & \cdots & -1 & 0 & \cdots & 0 & 1 \\ & & & 0 & & & & \end{bmatrix} \qquad (4-5)$$

式中 $\mathbf{0}$ 是一个 $H\times N$ 的零矩阵。在 ∂_z 的行中-1 和 1 之间有 $H-1$ 个零。因此 ∂_z 反映了垂向相邻块之间模型参数的差值。$N\times N$ 水平粗糙化矩阵 ∂_x 可表示为

$$\partial_x=\begin{bmatrix} \partial_{x_1} & & 0 & \\ & \partial_{x_2} & & \\ & & \ddots & \\ & 0 & & \partial_{xi} \end{bmatrix} \qquad (4-6)$$

式中 ∂_{xi} 是第 i 层的 $H\times H$ 水平粗糙化矩阵，可表示为

$$\partial_{x_i}=\begin{bmatrix} -w_i/h & w_i/h & & & \\ & -w_i/h & w_i/h & 0 & \\ & & \ddots & & \\ & 0 & & -w_i/h & w_i/h \\ & & & \cdots & 0 \end{bmatrix} \qquad (4-7)$$

因此 ∂_{xi} 表示了第 i 层侧向相邻单元之间模型参数的差异，对差值处置的基准是该层块的高度与宽度的比。

以视电阻率为例，如果有 n 个观测数据，表示为 d_j，$j=1$，2，…，n，σ_j 为每个数据的标准差,模型拟合数据的程度可利用拟合差标准定量化为

$$X^2=\sum_{j=1}^{n}\left(d_j-F_j[m]\right)^2 \Big/ \sigma_j^2 \qquad (4-8)$$

式中 X^2 是拟合差，$F[m]$ 是非线性正演函数作用于离散化模型 m 而产生的模型响应,对于给定数据集合 d 及误差,其正演问题的解可以用数学形式表述为：$d_j=F_j[m]$，$j=1$，2，…，n。数据模型表示为 $m\in E^N$，$d\in E^n$ 表示含有 n 个测量值的数据，F_j 是第 j 个数据的正演函数，拟合差 X^2 也可表示为：

$$X^2=\left\|\mathbf{W}d-\mathbf{W}F\left[m\right]\right\|^2 \tag{4—9}$$

$\mathbf{W}=diag\left\{1/\sigma_1,\ 1/\sigma_2,\ \cdots,\ 1/\sigma_n\right\}$ 为利用数据标准差进行归一化的矩阵：

$$\mathbf{W}=\begin{bmatrix}1/\sigma_1 & & & 0\\ & 1/\sigma_2 & & \\ & & \ddots & \\ 0 & & & 1/\sigma_n\end{bmatrix} \tag{4—10}$$

为了求解最小化问题，我们用拉格朗日乘子公式来平衡模型光滑和数据拟合程度，将反演的目标函数定义为：

$$U[m]=R_1+Q$$

$$=\|\partial_x m\|^2+\|\partial_z m\|^2+\mu^{-1}\{\|\mathbf{W}d-\mathbf{W}F[m]\|^2-X_*^2\} \tag{4—11}$$

式中 μ^{-1} 是拉格朗日乘子，X_*^2 为拟合差 X^2 的期望值。在实用中，规则化问题要预先设置拉格朗日乘子的值，然后求解拟合数据最好的模型。然而，这要求对模型粗糙度有事先的了解以及忽略数据过度拟合的可能性。因此，将求解的数据拟合到一个可接受误差范围内的模型会更合适。通常情况下，模型会随着允许误差的减小而变得粗糙。

对于精确的观测数据及初始模型 m_1, 可令 $X_*^2=0$ 将上述非线性问题线性化，对进行泰勒级数展开有如下公式

$$F\left[m_1+\Delta\right]=F\left[m_1\right]+J_1\Delta \tag{4—12}$$

式中 Δ 为模型修改量且 $\Delta=m_2-m_1$，m_2 为反演一次迭代的解，J_1 是初始模型 m_1 的雅可比矩阵（偏导数矩阵）：

$$J=\nabla_m F \tag{4—13}$$

其元素可表示为：$J_{ij}=\partial F_i\left[m\right]/\partial m_j$

代入公式（4—11），又有在 m_2 处的如下线性问题：

$$U=R_1+\mu^{-1}\left\{\left\|\mathbf{W}d_1-\mathbf{W}J_1m_2\right\|^2-X_*^2\right\} \tag{4—14}$$

式中 $d_1=d-F\left[m_1\right]+J_1m_1$ 是一种数据向量，因反演目标函数 U 在迭代过程中不断趋于极小值，若令 $\nabla\cdot u=0$（U 在 m 处的梯度），注意到此时 U 关于 $\mathbf{m}_2$ 是线性的，则有：

$$m_2=\left[\mu\partial^T\partial+\left(\mathbf{W}J_1\right)\mathbf{W}J_1\right]^{-1}\left(\mathbf{W}J_1\right)^T\mathbf{W}d_1 \qquad (4-15)$$

或

$$m_2=\left[\mu\left(\partial_x^T\partial_x+\partial_z^T\partial_z\right)+\left(\mathbf{W}J_1\right)^T\mathbf{W}J_1\right]^{-1}\left(\mathbf{W}J_1\right)^T\mathbf{W}d_1 \qquad (4-16)$$

由上述可以推出对模型m取微分寻找使U取极小值的第k次迭代公式：

$$m_{k+1}=\left[\mu\left(\partial_x^T\partial_x+\partial_z^T\partial_z\right)+\left(\mathbf{W}J_k\right)^T\mathbf{W}J_k\right]^{-1}\left(\mathbf{W}J_k\right)^T\mathbf{W}d_1 \qquad (4-17)$$

式中拉格朗日乘子μ为待求值，模型$m_{k+1}(\mu)$对求得的一系列μ值进行计算可以得到真正的拟合差

$$X_{k+1}(\mu)=\left\|\mathbf{W}d-\mathbf{W}F\left[m_{k+1}(\mu)\right]\right\| \qquad (4-18)$$

在反演的开始阶段，主要任务是减小拟合差，因为初始模型通常都与任何与观测数据相符合的模型相差较远，并且无论选取何μ值，$X_{k+1}(\mu)$都远大于$\boldsymbol{X}_*$。接下来是选取μ值以最小化$X_{k+1}(\mu)$，这可用一维线性搜索快速完成。迭代一定次数后，可选取μ使$X_{k+1}(\mu)$等于X_k。实际上，可能会有多个这样的μ值，此时可选其中最大的μ，因为它可使模型粗糙度最小。

4.3 求目标函数极小值的算法

按照 Tikhonov 和 Arsenin(1977)的思路解反演问题，取“正则化解”为对给定的λ、V和L最小化如下目标函数ψ的模型

$$\psi(m)=e^TV^{-1}e+\lambda m^TL^TL_m \qquad (4-19)$$

式中$e=d-F(m)$，d是数据向量，m是模型向量，F是正演模拟函数，e是残差向量。其中$d=\left[d^1\ \ d^2\ \ \cdots\ \ d^N\right]^T$的分量$d^i$为$\rho_{app}$（复数视电阻率）取对数后的振幅或其相位；$m=\left[m^1\ \ m^2\ \ \cdots\ \ m^M\right]^T$为地电模型的参数向量，可令$M$为模型单元个数，且每一$m^i$为取对数后的电阻率值（即$\log\rho$）。正则化参数$\lambda$是一个正数。正定矩阵$V$起残差向量$e$之协方差的作用，其是数据误差的一种度量。$\psi$的第二项是为了获得稳定的解而对地电模型的一种约束。本文中，我们将矩阵

L 取为简单的二阶微分算子，即 Lm 近似为 $\Delta(\log\rho)$。

在反演中，我们对观测数据和模型参数采用了其对数。这样做的意义在于：（1）保证了模型参数为正（即不允许有负的电阻率或电导率值）；（2）这是表示分成振幅（取对数）和相位的复数值的适宜方式（$\ln Z = \ln Z + i\theta$）；（3）取对数后的参数便于在迭代反演中模型参数的修正更大，对于电阻率有几个数量级变化的地电问题，这减少了达到一个可接受的解所需要的总迭代次数。

4.3.1 目标函数的梯度向量和海森矩阵

目标函数 $\psi(m)$ 的 M 维梯度向量 g 和 $M\times M$ 对称海森（Hessian）矩阵 H 定义为

$$g^j(m) = \partial_j \psi(m) \tag{4—20}$$

$$H^{jk}(m) = \partial_j \partial_k \psi(m),\quad j,\quad k = 1,\cdots,M \tag{4—21}$$

式中 ∂_j 表示对第 j 个函数变量的偏微分[$\psi(m)$ 即 $\psi(m^1, m^2, \cdots, m^M)$]。令 A 表示正演函数 F 的雅可比矩阵：

$$A^{ij}(m) = \partial_j F^i(m),\quad i = 1,\cdots,N;\quad j = 1,\cdots,M \tag{4—22}$$

由方程（4—6）有

$$g(m) = -2A(m)^T V^{-1}\left(d - F(m)\right) + 2\lambda L^T Lm \tag{4—23}$$

$$H(m) = 2A(m)^T V^{-1} A(m) + 2\lambda L^T L - 2\sum_{i=1}^{N} q^i B_i(m) \tag{4—24}$$

式中 B_i 是 F^i 的海森矩阵，$q = V^{-1}\left(d - F(m)\right)$。

我们也基于线性化 F 定义了一个近似的目标函数以及它的梯度向量和海森矩阵。对模型 m_{ref} 线性化，且定义

$$\tilde{F}(m; m_{ref}) = F(m_{ref}) + A(m_{ref})(m - m_{ref}) \tag{4—25}$$

$$\tilde{\psi}(m; m_{ref}) = \left(d - \tilde{F}(m; m_{ref})\right)^T V^{-1}\left(d - \tilde{F}(m; m_{ref})\right) + \lambda m^T L^T Lm \tag{4—26}$$

显然 $\tilde{\psi}$ 的梯度向量和海森矩阵可表示为

$$\tilde{g}(m; m_{ref}) = -2A(m_{ref})^T V^{-1}\left(d - \tilde{F}(m; m_{ref})\right) + 2\lambda L^T Lm \tag{4—27}$$

$$\tilde{H}\left(m_{ref}\right)=2A\left(m_{ref}\right)^{T}V^{-1}A\left(m_{ref}\right)+2\lambda L^{T}L \tag{4－28}$$

$\tilde{\psi}$ 是 m 的二次函数， $\tilde{g}$ 是 m 的线性函数，而 $\tilde{H}$ 则与 m 无关。事实上，

$$\begin{aligned}\tilde{\psi}\left(m;m_{ref}\right)&=\psi\left(m_{ref}\right)+g\left(m_{ref}\right)^{T}\left(m-m_{ref}\right)\\&\quad+\frac{1}{2}\left(m-m_{ref}\right)^{T}\tilde{H}\left(m_{ref}\right)\left(m-m_{ref}\right)\end{aligned} \tag{4－29}$$

$$\tilde{g}\left(m;m_{ref}\right)=g\left(m_{ref}\right)+\tilde{H}\left(m_{ref}\right)\left(m-m_{ref}\right) \tag{4－30}$$

易知 $\tilde{F}\left(m_{ref};m_{ref}\right)=F\left(m_{ref}\right)$ ， $\tilde{\psi}\left(m_{ref};m_{ref}\right)=\psi\left(m_{ref}\right)$ 以及 $\tilde{g}\left(m_{ref};m_{ref}\right)=g\left(m_{ref}\right)$ 而 $\tilde{H}\left(m_{ref}\right)$ 则是通过忽略方程（4－10）中的最后一项而得到的 $H\left(m_{ref}\right)$ 的一个近似。

在阐述最小化目标函数 $\psi\left(m\right)$ 的非线性共轭梯度算法之前，我们将首先介绍一下 Gauss-Newton 和 Mackie-Madden(1993)这两种最小化算法。在后文中，这些算法将相应的标示为 GN、MM 和 NLCG。每种算法产生一系列的模型 $m_0,m_1,\cdots$，且希望当 $l\to\infty$ 时 $\psi\left(m_l\right)\to\min_m\psi\left(m\right)$。

4.3.2 高斯－牛顿（Gauss-Newton）迭代法（GN）

Gauss-Newton 迭代可描述为递归的最小化 $\tilde{\psi}$ ，首先给出一个初始模型 m_0

$$m_0=given$$

它的模型序列需要满足下列方程式：

$$\tilde{\psi}\left(m_{l+1};m_l\right)=\min_m\tilde{\psi}\left(m;m_l\right),\qquad l=0,1,2,\cdots \tag{4－31}$$

为得到方程（4－31）的解，计算要求梯度向量 $\tilde{g}\left(m_{l+1};m_l\right)$ 为零。根据方程（4－30）， m_{l+1} 满足线性方程为

$$\tilde{H}_l\left(m_{l+1}-m_l\right)=-g_l \tag{4－32}$$

式中 $g_l\equiv g\left(m_l\right)$， $\tilde{H}_l\equiv\tilde{H}\left(m_l\right)$。假定 $\tilde{H}_l$ 非奇异，则 Gauss-Newton 迭代可写为

$$m_{l+1}=m_l-\left(\tilde{H}_l\right)^{-1}g_l \tag{4－33}$$

若采用 Levenberg-Marquardt 阻尼，即阻尼高斯－牛顿迭代法，它是马夸特（Marquardt）法中的一种简单形式，则方程（4－32）可写为

$$\left(\tilde{H}_l+\varepsilon_l I\right)\left(m_{l+1}-m_l\right)=-g_l \tag{4－34}$$

式中， I 是单位矩阵， ε_l 是一个允许随着迭代步长而变化的正的阻尼参数。

在 Gauss-Newton 算法中，先用高斯消元求出海森矩阵，然后用 Hassian $\tilde{H}_l$ 与 $-g_l$ 构成方程（4－33）右项，由方程（4－23）和（4－28）可知，要得到 g_l 和 $\tilde{H}_l$ 的值，那么就需要计算雅克比矩阵（Jacobi matrix）$A(m_l)$（雅克比矩阵的计算后面章节将有表述）。 Gauss-Newton 迭代法计算所需要的存储量主要是受到存储雅克比矩阵和海森矩阵的约束。在这种算法中，正演模拟计算和求目标函数 ψ 需要的存储量与观测数据数量和模型参数数量成线性关系，故对于大型的二维反演问题而言，其用于计算机计算的运算速度并不理想，如是高维的非线性问题会更难处理。

4.3.3 Mackie-Madden 算法（MM）

Mackie－Madden 算法是由 Mackie 和 Madden 于 1989 年首先提出的，并于 1993 年作了进一步的改进和实现。MM 算法是一种线性迭代反演算法，它是一个利用共轭梯度（CG）技术不完全求解线性方程组（4－32）的 Gauss-Newton 迭代过程。这种不完全性是指在经过 k 次的迭代后提前终止共轭梯度求解。因此，对每一次共轭梯度迭代 l，修改的模型参数 m_{l+1} 是由如下序列所产生的：

$$m_{l,0} = m_l$$

$$m_{l,k+1} = m_{l,k} + \alpha_{l,k} p_{l,k}, \qquad k = 0,1,\cdots,K-1$$

$$m_{l+1} = m_{l,k}$$

对每一次线性搜索 k ,向量 $p_{l,k}$ 是在模型空间的搜索方向，标量 $\alpha_{l,k}$ 是步长因子。我们简记

$$\tilde{g}_{l,k} \equiv \tilde{g}\left(m_{l,k}; m_l\right)$$

依照 CG 算法（Hestenes 和 Stiefel,1952），步长由如下公式给定

$$\alpha_{l,k} = -\frac{\tilde{g}_{l,k}^T p_{l,k}}{p^T{}_{l,k} \tilde{H}_l p_{l,k}} \qquad (4-35)$$

即求解一元的极小化问题且搜索方向的迭代式如下

$$\tilde{\psi}\left(m_{l,k} + \alpha_{l,k} p_{l,k}; m_l\right) = \min_{\alpha} \tilde{\psi}\left(m_{l,k} + \alpha p_{l,k}; m_l\right)$$

$$p_{l,0} = -C_l g_l$$

$$p_{l,k} = -C_l \tilde{g}_{l,k} + \beta_{l,k} p_{l,k-1} \quad , k = 1,2,\cdots,K-1 \qquad (4-36)$$

式中 $M \times M$ 正定矩阵 C_l 是一个已知的预处理因子，标量 $\beta_{l,k}$ 由下式给出

$$\beta_{l,k}=\frac{\tilde{g}_{l,k}^{T}C_l\tilde{g}_{l,k}}{\tilde{g}_{l,k-1}^{T}C_l\tilde{g}_{l,k-1}}$$

式（4－36）的第一项是预处理后的最速下降方向；第二项调整搜索方向以使其与前面的搜索方向共轭，即

$$p^{T}{}_{l,k}\tilde{H}_l p_{l,k'}=0 \quad ， \quad k'<k \tag{4－37}$$

共轭梯度算法的最后一部分是完成梯度向量的迭代：

$$\tilde{g}_{l,0}=g_l$$

$$\tilde{g}_{l,k+1}=\tilde{g}_{l,k}+\alpha_{l,k}\tilde{H}_l p_{l,k} \quad ， \quad k=0,1,\cdots,K-2$$

这是由式（4－30）得出的。

在 MM 算法中所需要的主要计算量是对正演函数 $F\left(m_l\right)$（用于计算 $\psi\left(m_l\right)$ 和 g_l）的求值以及对雅可比矩阵及其转置的运算。令

$$A_l\equiv A\left(m_l\right)$$

并定义

$$f_{l,k}=A_l p_{l,k} \quad ， \quad k=0,1,\cdots,K-1 \tag{4－38}$$

则式（4－35）的分母可写成

$$p_{l,k}^{T}\tilde{H}_l p_{l,k}=2f_{l,k}^{T}V^{-1}f_{l,k}+2\lambda p_{l,k}^{T}L^{T}Lp_{l,k}$$

并且对于梯度向量的迭代成为

$$\tilde{g}_{l,0}=-2A_l^{T}V^{-1}\left(d-F\left(m_l\right)\right)+2\lambda L^{T}Lm_l \tag{4－39}$$

$$\tilde{g}_{l,k+1}=\tilde{g}_{l,k}+2\alpha_{l,k}A_l^{T}V^{-1}f_{l,k}+2\alpha_{l,k}\lambda L^{T}Lp_{l,k} \quad ， \quad k=0,1,\cdots,K-2 \tag{4－40}$$

通过以上的计算变换，求解出雅克比矩阵和它的转置可以不需要计算它们的本身。但实际上这些演算中产生的替代矩阵是一个简单的拟正演问题的解，其计算量与实际的正演计算相当。MM 算法的优点是在每次 GN 迭代过程中进行了相对较少的 CG 迭代，这样做的主要目的是在需要计算全部雅克比矩阵的情况下保持了计算的高效性。对于大规模的非线性反演问题，MM 算法具有一定的优势。

4.4 非线性共轭梯度法（NLCG）

4.4.1 算法描述

在 MM 算法中，共轭梯度方法被用在 Gauss-Newton 型迭代中以不完全求解线性系统或者求目标函数的二次近似极小值，而非线性共轭梯度直接求解非二次的极小值问题，突破了线性化迭代反演的框架。NLCG 算法采用非线性共轭梯度的 Polak-Ribiere 变量以最小化式（4－19）的目标函数ψ。

非线性共轭梯度的模型序列由一系列的一元极小化或沿着计算的搜索方向的线性搜索所确定：

$m_0 = given$

$$\psi\left(m_l + \alpha_l p_l\right) = \min_{\alpha} \psi\left(m_l + \alpha p_l\right) \qquad (4-41)$$

$m_{l+1} = m_l + \alpha_l p_l$ ，$l = 0,1,2,\cdots$

p 是搜索的方向，α_l 是步长因子，l 是迭代的次数，搜索方向是对线性共轭梯度的近似，其迭代格式为

$p_0 = -C_0 g_0$

$$p_l = -C_l g_l + \beta_l p_{l-1} \quad ，\ l = 1,2,\cdots \qquad (4-42)$$

式中，在 Polak-Ribiere 技术中，

$$\beta_l = \frac{g_l^T C_l \left(g_l - g_{l-1}\right)}{g_{l-1}^T C_{l-1} g_{l-1}}$$

$-C_l g_l$ 是最快速的下降方向。它是通过在 m_l 处直接求目标函数ψ的导数的极小值来获得的，与线性共轭梯度法不同的是其搜索方向不必关于某个固定矩阵共轭，但要满足条件：

$$p_l^T \left(g_l - g_{l-1}\right) = 0 \quad ，\ l > 0 \qquad (4-43)$$

式（4－41）的极小值问题不是二次的，且需要某种迭代技术来求解。因为它只有一个未知量，可尝试将其作为一个全局最优化问题来处理，也就是寻找ψ关于α的一个全局极小值。这样做所获得的结果就优于没有区分局部极小还是全局极小的 Gauss-Newton 算法。然而，全局最优化算法可能导致在每一步非线性共轭梯度迭代中需要多次计算正演问题。考虑到 MT 正演问题的计算量，NLCG

算法并不将式(4－41)做全局最小化而是将其处理成计算量很小(高永才，2004)。

在本算法中，线性搜索算法对 Gauss-Newton 方法作了一元形式的改进。为方便起见，用Φ_l表示要极小化的一元函数，其 Gauss-Newton 近似为$\tilde{\Phi}_l$：

$$\begin{aligned}&\Phi_l(\alpha)\equiv\psi(m_l+\alpha p_l)\\&\tilde{\Phi}_l(\alpha;m_{ref})\equiv\tilde{\psi}(m_l+\alpha p_l;m_{ref})\end{aligned}\qquad(4-44)$$

线性搜索产生的模型序列为

$m_{l,k}=m_l+\alpha_{l,k}p_l$ ， $k=0,1,2,\cdots$

式中

$$\begin{aligned}&\alpha_{l,0}=0\\&\tilde{\Phi}_l(\alpha_{l,k+1};m_{l,k})=\min_{\alpha}\tilde{\Phi}_l(\alpha;m_{l,k})\ ，\ k=0,1,2,\cdots\end{aligned}\qquad(4-45)$$

因为$\tilde{\Psi}(m;m_{l,k})$是m的二次函数，$\tilde{\Phi}_l(\alpha;m_{l,k})$是$\alpha$的二次函数，式（4－45）可由下式求解它的极小值

$$\alpha_{l,k+1}=\alpha_{l,k}-\frac{g_{l,k}^T p_l}{p_l^T\tilde{H}_{l,k}p_l}\qquad(4-46)$$

其中

$g_{l,k}\equiv g(m_{l,k})$

$\tilde{H}_{l,k}\equiv\tilde{H}(m_{l,k})$

非线性共轭梯度法对高斯－牛顿算法的改进有如下几点：

（1）在线性搜索中找到的最优的ψ模型，将其表示为$m_{l,best}\equiv m_l+\alpha_{l,best}p_l$

（2）如在迭代过程中Φ_l增大（$\Phi_l(\alpha_{l,k})>\Phi_l(\alpha_{l,k-1})$），则用下式计算下一个步长

$$\alpha_{l,k+1}=\tfrac{1}{2}(\alpha_{l,k}+\alpha_{l,best})\qquad(4-47)$$

（3）在线性搜索的第二步或以后的步骤中，如果当前的模型和前面的最优模型包含一个如下意义的极小值

$$\Phi_l'(\alpha_{l,best})\Phi_l'(\alpha_{l,k})<0$$

则$\alpha_{l,k+1}$不再由式（4－46）计算，而是取为$\Phi_l(\alpha)$三次近似函数的局部极小点。该三次近似函数与Φ_l及其在$\alpha=\alpha_{l,k}$和$\alpha=\alpha_{l,best}$的导数相符合。

当由 $\Phi_l(\alpha)$ 的二次或三次近似函数所预测的 $\alpha_{l,k+1}$，其对应的目标函数值 $\Phi_l(\alpha_{l,k+1})$ 在指定的容许误差范围内时，则认为该线性搜索是收敛的。通常在 Gauss-Newton 更新中，收敛条件为

$$\left|\Phi\left(\alpha_{l,k+1}\right)-\tilde{\Phi}_l\left(\alpha_{l,k+1};m_{l,k}\right)\right|\le\tau\Phi_l\left(\alpha_{l,k+1}\right)$$

式中 $\tau<<1$ 为容许误差。如果在指定的最大搜索次数中没有收敛或出现 $\Phi_l(\alpha_{l,k+1})>1.5\Phi_l(\alpha_{l,best})$ 则认为该线性搜索是失败的。无论如何，第 l 次线性搜索的最终结果取为所发现的最优模型：

$$m_{l+1}=m_{l,best}$$

如果线性搜索收敛，则新的搜索方向 p_{l+1} 由式（4－42）计算。如果失败，p_{l+1} 取为最速下降方向式（4－42）的第一项，打破和前面搜索方向的共轭性。在这种情况下，即使目标函数 ψ 和它的梯度在 $m_{l,k}$ 处的值有很大的跳动，线性搜索的终止条件也将被满足，而且 $m_{l,k+1}$ 将是线性搜索的准确结果。线性搜索结束时不需要任何附加条件，除非需要更新额外的模型 $\left(m_{l,k+2}\right)$ 或者是在新模型处需要计算二次导数。这样，当非线性共轭梯度迭代达到这一点时，这里的正演函数 F 在所有的搜索方向都扮演着线性角色，每个极小值仅仅需要一次迭代，并且剩下的计算与 MM 算法中的共轭梯度计算基本相同。实践证明，NLCG 反演算法在进行大规模的非线性反演问题求解时，对二维大地电磁资料的处理是比较有效的，既有较高的分辨率，又有较快的计算速度和较小的存储量。

4.4.2 NLCG 反演的流程图

NLCG 算法产生一个双下标的模型系列 $m_{l,k}$，其中 l 标示共轭梯度循环，k 标示 Gauss-Newton 迭代。下面我们简述了一个具有单步线性搜索的 NLCG 反演的算法结构（见图 4-1），输入为初始模型 m_0：

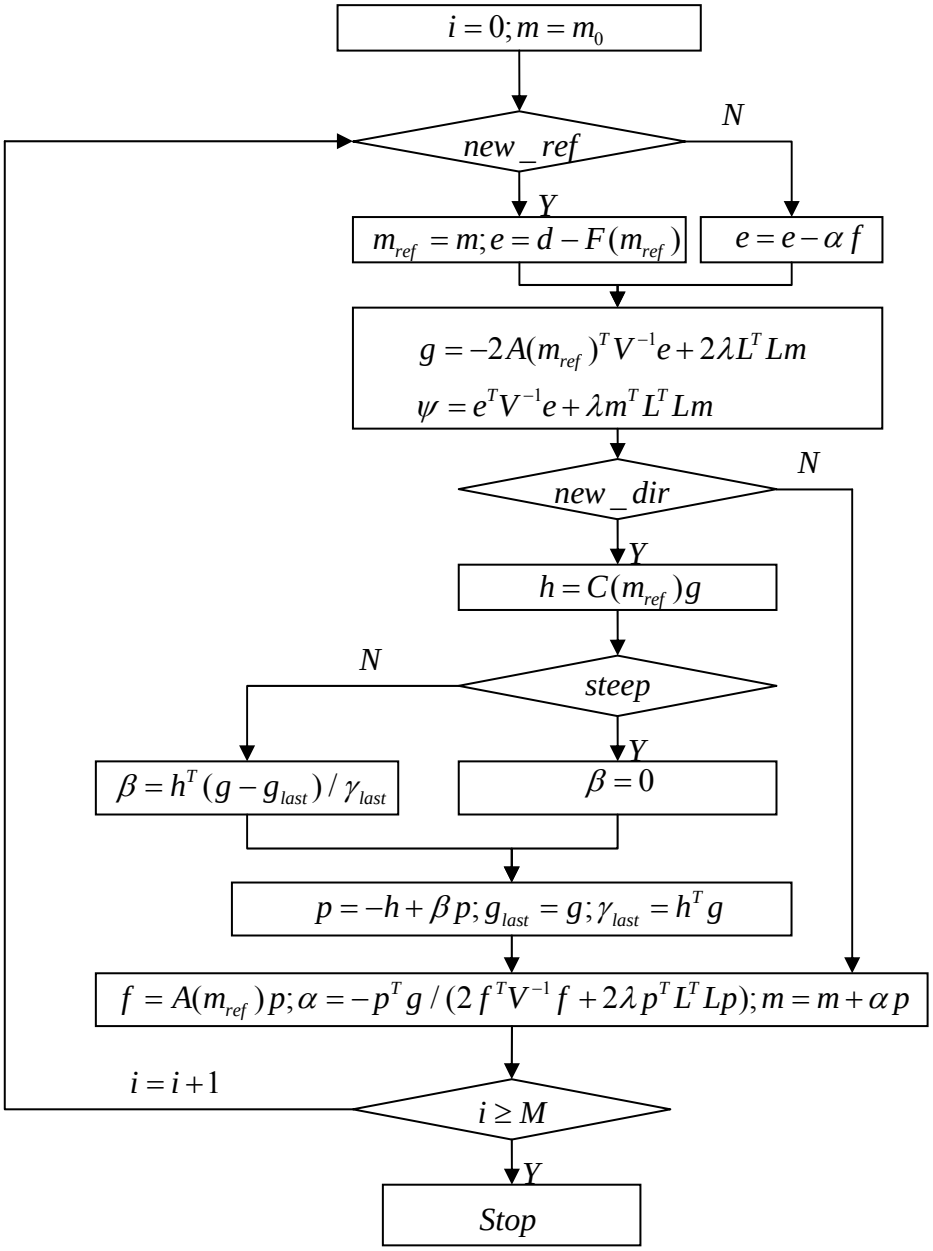


图 4-1 NLCG 算法流程

上述计算α的公式请参见式（4－46）。算法描述中的三个标识*new_ref*，*new_dir* 和*steep* 的含义为：如果当前模型用作线性化的参考模型则*new_ref* 置为 1（true）；如果改变了搜索方向则*new_dir* 为 1；如果最新计算的搜索方向被设为最速下降方向则*steep* 为 1，因此也就打破了式（4－44）的共轭性条件。在反演中，这三个标识均被初始化为 1。

4.4.3 搜索方向的预处理因子

在 MM 算法和 NLCG 算法的共轭梯度迭代中都使用了预处理因子 C_l。它可能会对共轭梯度的效率有很大影响。对其选取要考虑两个方面：使用预处理因子的计算量大小以及其使梯度向量转变成为有效的搜索方向的能力。

本文中预处理因子采用了如下形式（Rodi，W.and Mackie，R.L，2001）：

$$C_l = \left(\gamma_l I + \lambda L^T L\right)^{-1} \qquad (4-48)$$

式中 γ_l 是已知的标量，I 是单位矩阵。在计算 h 时，我们通过将预处理因子作用于向量 g 求解下面的线性方程组

$$\left(\gamma_l I + \lambda L^T L\right) h = g$$

可用线性共轭梯度方法求解该方程组。

方程（4－48）的基本原理是对于一个存在的有效算子，其在某种意义下起到近似的 Hessian $\tilde{H}_l$ 的逆。应用 C_l 的效率来自上述线性系统中关于 h 的简易性和稀疏性。求解该方程组所需的计算量小于一次正演函数计算的求值。因此这对 NLCG 反演只添加了很小的计算量，而且近似海森矩阵的逆可以从预处理因子的第二项导出。

4.5 灵敏度分析

研究大地电磁测深资料的灵敏度，旨在认识与了解它的变化特征和规律，为更有效地进行反演解释提供依据。从理论上说，灵敏度高的资料更适合作为反演解释的参数。虽然灵敏度与分辨能力是两个不同的概念，但就模型对观测资料的作用而言是有某种联系的。

在 Madden（1972）和 Tripp，Honhamnn & Swift（1984）发展的一种基于互易原理的方法基础上，Madden（1990）和 Mackie（1991）详细地讨论了大地电磁问题的灵敏度分析和互易关系。这在 NLCG 反演中也是一个很重要的问题。在 MT 中，观测资料通常表示成频率域的视电阻率 ρ_T 和阻抗相位 ϕ 形式。

$$\rho_T = \frac{1}{\omega\mu_0}\left|\frac{E}{H}\right|^2 \quad , \quad \phi = phase\left(\frac{E}{H}\right) \qquad (4-49)$$

在二维的情况下，式（4－49）有 TE 和 TM 之分，视电阻率 ρ_T 表示为复视

电阻率 ρ_{app}。由于在反演过程中，观测资料是视电阻率和阻抗相位，因此需要求解的是视电阻率和阻抗相位的灵敏度而不是电场和磁场本身的灵敏度。如果我们要获得关于磁化率参数的灵敏度，对方程（4－49）求关于第 l 个磁化率参数的偏导数：

$$\begin{aligned}\frac{\partial \rho_{app}}{\partial \kappa_l} &= \frac{\partial}{\partial \kappa_l}\left(\frac{1}{\omega\mu_0}\left|\frac{E}{H}\right|^2\right) \\ &= \frac{2}{\omega\mu_0}\left|\frac{E}{H}\right|\left[\frac{1}{|H|}\frac{\partial|E|}{\partial\kappa_l}-\frac{|E|}{|H|^2}\frac{\partial|H|}{\partial\kappa_l}\right]\end{aligned} \tag{4－50}$$

电场 E 在频率域中表示为一个复数，如下

$$E = |E|\exp\left(i\Phi_E\right) \tag{4－51}$$

E 是电场的幅值，Φ_E 为电场信号的相位，E 和 Φ_E 都作为模型参数的函数，对方程（4－51）求磁化率 κ_l 的偏导数，

$$\begin{aligned}\frac{\partial E}{\partial \kappa_l} &= |E|\frac{\partial}{\partial\kappa_l}\left[\exp\left(i\Phi_E\right)\right]+\frac{\partial|E|}{\partial\kappa_l}\exp\left(i\Phi_E\right) \\ &= \frac{E}{|E|}\frac{\partial|E|}{\partial\kappa_l}+i|E|\exp\left(i\Phi_E\right)\frac{\partial\Phi_E}{\partial\kappa_l}\end{aligned} \tag{4－52}$$

可以推出，$\dfrac{1}{E}\dfrac{\partial E}{\partial\kappa_l}=\dfrac{1}{|E|}\dfrac{\partial|E|}{\partial\kappa_l}+i\dfrac{\partial\Phi_E}{\partial\kappa_l}$ （4－53）

上式的实部和虚部分别为，

$$\frac{\partial|E|}{\partial\kappa_l}=|E|\mathrm{Re}\left(\frac{1}{E}\frac{\partial E}{\partial\kappa_l}\right) \tag{4－54}$$

$$\frac{\partial\Phi_E}{\partial\kappa_l}=\mathrm{Im}\left(\frac{1}{E}\frac{\partial E}{\partial\kappa_l}\right) \tag{4－55}$$

同样的在磁场中，有

$$\frac{\partial|H|}{\partial\kappa_l}=|H|\mathrm{Re}\left(\frac{1}{H}\frac{\partial H}{\partial\kappa_l}\right) \tag{4－56}$$

和

$$\frac{\partial\Phi_H}{\partial\kappa_l}=\mathrm{Im}\left(\frac{1}{H}\frac{\partial H}{\partial\kappa_l}\right) \tag{4－57}$$

将式（4－54）和式（4－56）带入方程（4－50），可得到视电阻率关于磁化率参数的灵敏度表达式：

$$\begin{aligned}\frac{\partial\rho_{app}}{\partial\kappa_l}&=\frac{2}{\omega\mu_0}\left|\frac{E}{H}\right|\left[\frac{|E|}{|H|}\mathrm{Re}\left(\frac{1}{E}\frac{\partial E}{\partial\kappa_l}\right)-\frac{|E|}{|H|}\mathrm{Re}\left(\frac{1}{H}\frac{\partial H}{\partial\kappa_l}\right)\right]\\&=2\rho_{app}\left[\mathrm{Re}\left(\frac{1}{E}\frac{\partial E}{\partial\kappa_l}\right)-\mathrm{Re}\left(\frac{1}{H}\frac{\partial H}{\partial\kappa_l}\right)\right]\end{aligned}\tag{4－58}$$

在已知视电阻率的情况下，可利用上式计算出视电阻率关于磁化率参数的灵敏度，并且$\partial E/\partial\kappa_l$和$\partial H/\partial\kappa_l$可以用前述方法计算出来。

就阻抗相位关于磁化率参数的灵敏度表达式，对于如下比值

$$\frac{E}{H}=\frac{|E|\exp(i\Phi_E)}{|H|\exp(i\Phi_H)}=\left|\frac{E}{H}\right|\exp\left[i(\Phi_E-\Phi_H)\right]\tag{4－59}$$

定义阻抗相位为：$\phi=\Phi_E-\Phi_H$　　（4－60）

由式（4－55）和式（4－57）可得，

$$\frac{\partial\phi}{\partial\kappa_l}=\frac{\partial\Phi_E}{\partial\kappa_l}-\frac{\partial\Phi_H}{\partial\kappa_l}=\mathrm{Im}\left(\frac{1}{E}\frac{\partial E}{\partial\kappa_l}\right)-\mathrm{Im}\left(\frac{1}{H}\frac{\partial H}{\partial\kappa_l}\right)\tag{4－61}$$

最后，可用方程（4－58）和（4－61）计算视电阻率和阻抗相位关于磁化率参数的灵敏度。

同理可推导出视电阻率和阻抗相位关于电阻率参数的灵敏度，从方程（4－49）开始求关于电阻率参数的偏导数，得到

$$\begin{aligned}\frac{\partial\rho_{app}}{\partial\rho_l}&=\frac{\partial}{\partial\rho_l}\left(\frac{1}{\omega\mu_0}\left|\frac{E}{H}\right|^2\right)\\&=\frac{2}{\omega\mu_0}\left|\frac{E}{H}\right|\left[\frac{1}{|H|}\frac{\partial|E|}{\partial\rho_l}-\frac{|E|}{|H|^2}\frac{\partial|H|}{\partial\rho_l}\right]\end{aligned}\tag{4－62}$$

分别用式（4－54）、（4－56）和式（4－55）、（4－57）可得到视电阻率和阻抗相位关于电阻率参数的灵敏度表达式：

$$\frac{\partial\rho_{app}}{\partial\rho_l}=2\rho_{app}\left[\mathrm{Re}\left(\frac{1}{E}\frac{\partial E}{\partial\rho_l}\right)-\mathrm{Re}\left(\frac{1}{H}\frac{\partial H}{\partial\rho_l}\right)\right]\tag{4－63}$$

和

$$\frac{\partial\phi}{\partial\rho_l}=\frac{\partial\Phi_E}{\partial\rho_l}-\frac{\partial\Phi_H}{\partial\rho_l}=\mathrm{Im}\left(\frac{1}{E}\frac{\partial E}{\partial\rho_l}\right)-\mathrm{Im}\left(\frac{1}{H}\frac{\partial H}{\partial\rho_l}\right)\tag{4－64}$$

以上是视电阻率和阻抗相位关于模型参数的灵敏度表达式。

4.6 雅可比矩阵与向量乘积的整体计算

我们以 TE 模式为例，讨论雅可比矩阵 A 以及 A_x 与 $A^T y$ 的计算问题。灵敏度矩阵描述了介质物理参数的扰动所引起的地表观测结果的扰动。虽然没有确切地指出要用哪个地表观测结果反演，但它们都涉及到地表电磁场的导数，以及包括象 $\partial E_x/\partial\sigma$，$\partial H_y/\partial\sigma$，$\partial H_x/\partial\sigma$，$\partial E_y/\partial\sigma$ 的项。由复数视电阻率的定义

$$\rho_{app}=\frac{i}{\omega\mu}\left(\frac{E}{H}\right)^2 \tag{4－65}$$

可知，雅可比矩阵的元素可表示为

$$A_{ij}=\frac{\partial\rho_{appi}}{\partial\sigma_j}=\frac{2i}{\omega\mu}\left[\frac{E_i}{H_i^2}\frac{\partial E_i}{\partial\sigma_j}-\frac{E_i^2}{H_i^3}\frac{\partial H_i}{\partial\sigma_j}\right]\sim\frac{\partial\left(E_i,H_i\right)}{\partial\sigma_j} \tag{4－66}$$

由此，若要实际地构造灵敏度矩阵，就应在每一地表测点位置放一单位点源，然后对每一频率计算由该点源所引起的介质内部的场，这些场乘以原来的电场 $E_x(s)$ 就会得到 $\partial E/\partial\sigma$ 和 $\partial H/\partial\sigma$ 的值，即

$$\partial(E,H)/\partial\sigma=G(r,s)\cdot E_x(s)\cdot(area)$$

而后则可计算所需要的灵敏度元素。

然而在 NLCG 反演中，我们不必构造灵敏度矩阵。注意到灵敏度矩阵 **A** 乘以任一个向量 **x** 就是如下形式：

$$p=Ax\sim\sum_j\frac{\partial\left(E,H\right)}{\partial\sigma_j}x_j \tag{4－67}$$

j 表示模型参数向量 p 的每一分量都定义在某一测点位置的某一频率。我们可利用线性迭加原理在没有计算灵敏度矩阵 A 的情况下计算向量 p。迭加原理表明如果 T 为线性系统，则对于任意两个输入 l_1 和 l_2 以及任意的标量常数 C 有

$$\begin{cases}T\left[l_1+l_2\right]=T\left[l_1\right]+T\left[l_2\right]\\T\left[Cl_1\right]=CT\left[l_1\right]\end{cases} \tag{4－68}$$

可知，上述关系适用于任意个输入量的情形。在我们的情况下，线性系统 T 表示麦克斯韦方程，且输入为电场源或磁场源。由于地表场的灵敏度即是由介质中的单位电场源乘以该处的电场所引起的响应，因此向量 p 便为

$$T\left[J_0(1)E_0(1)\right]x(1)+T\left[J_0(2)E_0(2)\right]x(2)+\cdots+T\left[J_0(n)E_0(n)\right]x(n) \quad (4-69)$$

式中 $J_0(i)$ 为单位电场源，$E_0(i)$ 为原来的电场，$x(i)$ 为向量 x 的分量，且 i 表示模型单元。然而，利用迭加原理，我们可将向量 p 另外表示为

$$T\left[J_0(1)E_0(1)x(1)+J_0(2)E_0(2)x(2)+\cdots+J_0(n)E_0(n)x(n)\right] \quad (4-70)$$

显然，这等价于同时将所有的源放在介质中且计算一次正演以确定在地表的影响。这样计算出的地表场给出了在某一测点、某一频率灵敏度矩阵乘以向量 x 所要求的结果，且这样做并没有实际地计算灵敏度矩阵。由此可见，采用本方法在计算时间方面我们有了极大的改进。由常规方法计算灵敏度矩阵，要求对于每一频率、每一模型参数做一次正演。利用互易关系，对每一频率可将解正演问题的次数减少到地表测点的个数。然而，我们计算 A_x 的方法对每一频率仅需做一次正演。

同样，可用类似的方法来计算 A^T 乘以任一向量 y。令向量 $Q=A^Ty$，则有

$$Q\sim\sum_{freq}\sum_{surf}\frac{\partial(E,H)_{surf}}{\partial\sigma}y^*_{surf} \quad (4-71)$$

如前，我们可利用线性迭加原理计算向量 Q。由互易性，可知地表场的灵敏度是由地表的单位电场源或磁场源（这取决于求解的是电场扰动或磁场扰动）所引起的响应。因此，向量 Q 即是

$$\left\{T\left[s(1)\right]y^*(1)+T\left[s(2)\right]y^*(2)+\cdots+T\left[s(k)\right]y^*(k)\right\}E_0(i) \quad (4-72)$$

式中 $s(k)$ 表示地表的单位电场源或磁场源，$y(k)$ 为向量 y 的分量，k 表示地表位置。同前，$E_0(i)$ 为第 i 个模型单元原来的电场。由线性迭加原理，可将向量 Q 重写为

$$\left\{T\left[s(1)y^*(1)+s(2)y^*(2)+\cdots+s(k)y^*(k)\right]\right\}E_0(i) \quad (4-73)$$

这对应于所有的源同时分布在地表的一次正演模拟。

因此，我们能够在没有精确地计算灵敏度矩阵的情况下，通过对每个频率求解源分布在介质中或地表的“伪正演问题”来整体计算 A_x 与 A^Ty 的值,从而极大的减小了计算工作量，提高了反演速度。

通过实验，我们知道 OCCAM 反演和 NLCG 反演的迭代过程，都稳定收敛，

它们的不同之处在于：（1）OCCAM 反演对初始模型依赖程度小，成像效果良好；而 NLCG 反演则对初始模型依赖性较大，成像效果的优劣与给定的初始模型条件好坏密切相关；（2）OCCAM 反演，由于其需要直接计算雅可比矩阵，每次迭代要调用很多次正演函数，因此反演速度非常慢；而 NLCG 反演，由于其利用了线性系统（即麦克斯韦方程组）的迭加原理和格林函数（点源场）的性质（特别是互易关系），策略性的实现了雅可比矩阵与一个向量乘积的整体计算，从而极大的提高了反演速度。综合这些分析可知，OCCAM 反演和 NLCG 反演具有很好的互补性：在对初始模型的依赖性方面，OCCAM 反演的优势可以弥补 NLCG 反演的缺点；而在反演速度方面，NLCG 反演的优势则能弥补 OCCAM 反演的缺点。正是基于这些认识，我们对非线性共轭梯度的反演方案提出了一种改进措施，即首先以 OCCAM 一维反演形成拟二维地电剖面，然后再用 NLCG 对此初始模型做二维反演。尽管简单，但实践证明这种改进方案非常有效。

通过数值模拟实验表明，我们所提出的改进方案（即 OCCAM 一维反演和 NLCG 二维反演的联合运用）取得了良好的反演效果，这种效果取决于两个方面：（1）OCCAM 一维反演能够为二维反演提供较理想的初始模型；（2）NLCG 反演是一种成像效果较好的二维反演方法。因此，只要上述两个条件成立，我们的反演方案就能取得好的成像效果。当然，如果两个条件之一不再成立，则反演方法将不在适用，这也是本反演方案的局限性。

4.7 小结

由于大地电磁测深法的体积效应，我们在实际工作中获得的观测资料仅是对测点以下一定范围内的真实地电信息的空间加权平均的一个特定采样。因此，大地电磁观测资料所反映的地电信息只是真实电性结构的某种平滑后的图像。有鉴于此，从单独的某个频点的视电阻率值很难判定它是代表哪个深度的信息，同时观测资料对构造电性参数变化的灵敏度也受到限制。这也是大地电磁资料反演存在多解性的物理原因。在均匀半空间中，可以通过解析式来分析 MT 测深资料的灵敏度函数的变化特征，这对于分析大地电磁资料的灵敏度函数具有指导意义。在二维大地电磁反演问题中，我们推导了视电阻率和阻抗相位分别关于磁化率和电阻率参数的灵敏度，可将此推导的公式应用于反演程序中以实现灵敏度矩阵的计算。