地球物理反演

（讲义）

主讲：朱良保

符号规则：黑体小写拉丁字母代表矢量，黑体大写拉丁字母代表矩阵（非一维）。
带下角标的非黑体拉丁字母代表分量。

1．前言

物理科学中一个非常重要的研究领域就是如何由观测数据来推断物理参数。如：太阳的内部结构，储油层的深度，Moho面的深度，核幔边界的形态等。如果给定物理体系的参数，一般来说，由物理定律能够计算出与观测数据相对比的理论数据。由物理定律根据给定的物理参数计算出数据是正演问题。如图1

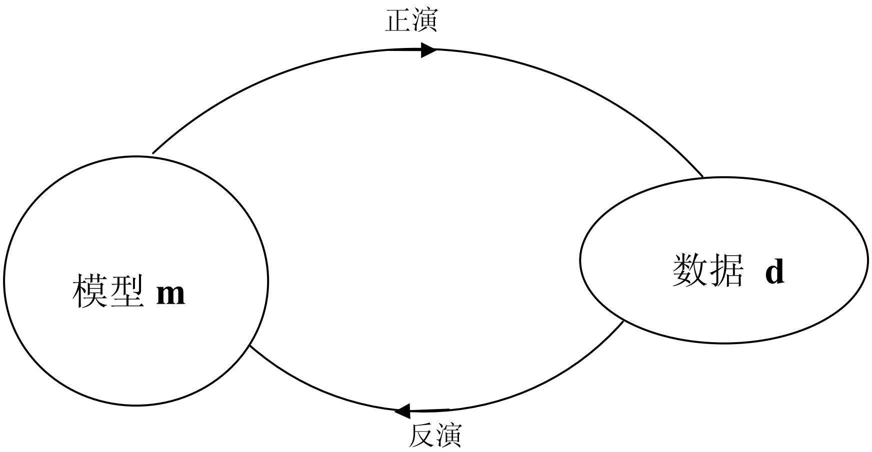

图 1. 正反演的传统定义

反演问题是根据一组观测数据来重建物理模型。需要强调的是，任何形式的反演过程都必须借助正演手段。没有理论上的正演，就不可能把观测数据有效地与物理参数联系起来，反演就失去方向。

对于有些问题，在理想的情况下，运用反演理论可以唯一地求出物理模型。如地震学中的Herglotz-Wiechert 反演理论，应用理想的走时数据，假定地层的速度是一维的并随深度单调增加，可以唯一的反演出地层的速度结构。在这个理论中，有两条很重要的假设，其一是数据是无误差的，其二是速度是一维并单调增加的。这两条在实际过程中都不可能达到。尽管非线性反演方案在数学上看起来很漂亮，但在实际运用中却非常有限。首先，准确的反演技术一般只能运用于理想的情况，但在实际中很难满足反演条件，如Herglotz-Wiechert反演理论。另外，准确的反演技术在反演中经常出现不稳定。更重要的一点是，实际物理模型一般是连续的。也就是说，模型的自由度是无穷维的。而实际过程中，观测数据是有限的并且是有误差的。有限的带有误差的数据不足于保证反演的唯一性。Backus and Gilbert (1967，

1968)证明了，在线性反演中由于数据量的不足以及误差，反演的不唯一性是必然的。非线性反演更是如此。

反演的非唯一性特征意味着，存在许多反演模型能够解释观测数据，由观测数据反演得到的模型不一定就是真实的模型。由图1表征的反演过程过于简单了。我们必须做一些其他的事情。实际的反演过程分两步进行。假如用$\mathbf{m}$表示真实的模型，$\mathbf{d}$表示数据。第一步由数据反演推断出模型$\tilde{\mathbf{m}}$，这一步为推断。既然有许多模型都能解释拟合观测资料，那么推断出的模型到底与真实模型存在什么样的关系，多大程度上表征了真实的模型，推断出的模型的误差到底有多大, 这些问题我们必须回答。没有模型误差分析以及不讨论模型的分辨率，反演出的模型就没有太大的意义。于是乎我们必须作第二步工作，评估模型。所以实际的反演过程可以表征为：反演=推断+评估。图2是实际反演过程的示意图。

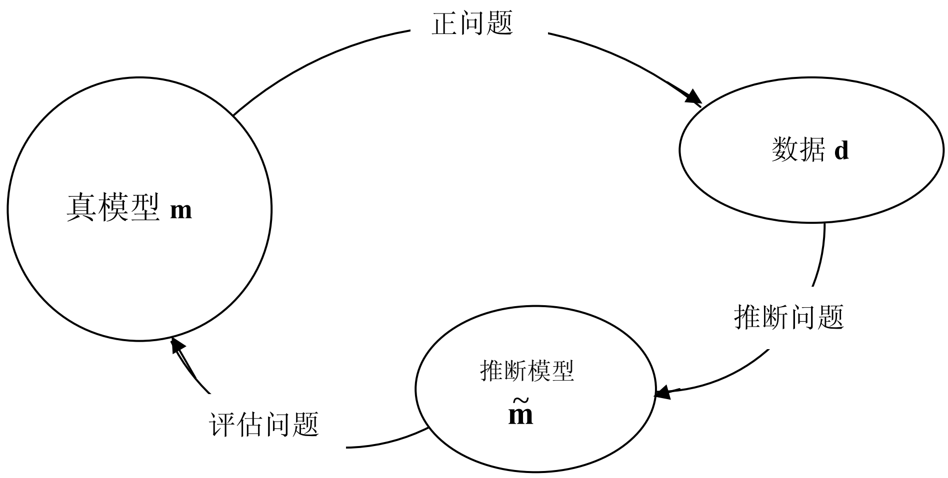

图 2.反演过程是推断加评估

一般来说，观测数据是离散的。而模型可以是离散的，也可以是连续的。就模型而言，反演分离散方法和连续方法，处理上有所不同。模型参数与数据的关系有时是线性的，有时是非线性的。对于线性问题，目前的有比较成熟的解决方案。非线性问题比较复杂，还没有找到很好的方案解决模型评估问题。一种方案是对非线性问题做局部近似使其线性化，然后采取循环迭代，逐步接近非线性问题的解，其结果依赖于初始模型的选择。另一种方案是模型空间的全局搜索。目前，无论哪种方案都不能很好地解决非线性反演问题。非线性方法的研究是一个挑战。

2．离散线性反演

如果模型可用有限个参数来描述，则反演问题称为离散型反演。设模型矢量为 $\mathbf{m}=(m_1,m_2,\cdots,m_n)^T$，数据矢量为$\mathbf{d}=(d_1,d_2,\cdots,d_m)^T$，连接数据与模型的矩阵为$\mathbf{A}$，称为理论算子。因为它包含了正演计算中所有的物理数学信息。$\mathbf{A}$作用到模型矢量上为$\mathbf{Am}$。在实际观测中总会有误差，数据的误差用误差矢量$\boldsymbol{\varepsilon}$表示。模型矢量与数据的关系可以表示为

$$\mathbf{d}=\mathbf{Am}+\boldsymbol{\varepsilon} \qquad (1)$$

模型参数的选择带有的一定主观性任意性。例如，描述地球的密度你也许选择地幔是均

匀的，地核也是均匀的，那么密度模型就可用两个参数表示。如果想更精确一些描述，你可以用球谐函数在球面上展开，深度上用多项式来表示，这样就需更多的参数来描述模型。有时对同一个反演问题，不同的人根据自己的喜好用不同的参数来描述模型。由于这种任意性，用模型参数描述的模型就不一定是真实的模型。毕竟真实的模型我们是不知道的。不管用哪种模型参数来描述，总有它的合理性，并对众多模型的选择是一种约束。我们不妨把 $\mathbf{m}$ 叫做理论真模型，或简称真模型。由数据反演出的模型叫估计模型，用 $\tilde{\mathbf{m}}$ 表示。由于数据的误差及某些我们不清楚的因素，一般来说估计模型不会等于真模型。求解方程（1）的方法很多［如，*Menke*, 1984;*Tarantola*, 1987; *Parker*, 1994]，归根到底就是求逆算子，使得

$$\tilde{\mathbf{m}} = \mathbf{A}^{-g}\mathbf{d} \qquad (2)$$

算子 $\mathbf{A}^{-g}$ 称为矩阵 $\mathbf{A}$ 的广义逆。一般来说，数据个数与模型的个数是不相同的，所以 $\mathbf{A}$ 不是方阵，也就不存在逆矩阵。但广义逆是能够给出的。以后再讨论广义逆怎么求。由（1）和（2）可以得出估计模型与真模型之间的关系。把（1）代进（2）得

$$\tilde{\mathbf{m}} = \mathbf{A}^{-g}\mathbf{A}\mathbf{m} + \mathbf{A}^{-g}\boldsymbol{\varepsilon} \qquad (3)$$

矩阵 $\mathbf{A}^{-g}\mathbf{A}$ 称为分辨矩阵或分辨核。它是一个算子，表为

$$\mathbf{R} \equiv \mathbf{A}^{-g}\mathbf{A} \qquad (4)$$

如果 $\mathbf{R} = \mathbf{I}$，则 $\tilde{\mathbf{m}}$ 的分量与 $\mathbf{m}$ 的分量一一对应，估计模型完全分辨。否则，估计模型就是真模型的线性组合。模型的分辨由数据量和数据的结构决定，而与数据的质量无关。（3）式右边的第二项表示数据的误差对估计模型的影响。虽然我们不知道数据误差的细节，否则我们就可以把他们从数据中消除掉，但通过统计分析可以得到数据误差引起的模型误差。设数据的各分量间相互独立，第 i 个分量的标准误差为 σ_{d_i}。由（3）式得

$$\Delta\tilde{\mathbf{m}} = \mathbf{A}^{-g}\Delta\varepsilon$$

由（1）式知 $\qquad \Delta\mathbf{d} = \Delta\varepsilon$

$$\therefore \qquad \Delta\tilde{\mathbf{m}} = \mathbf{A}^{-g}\Delta\mathbf{d}$$

估计模型的协方差矩阵为 $\quad \left\langle \Delta\tilde{\mathbf{m}}(\Delta\tilde{\mathbf{m}})^T \right\rangle = \mathbf{A}^{-g}\left\langle \Delta\mathbf{d}(\Delta\mathbf{d})^T \right\rangle \left(\mathbf{A}^{-g}\right)^T$

则第 α 个模型分量的方差为

$$\sigma_{m_\alpha}^2 = (\mathbf{A}^{-g})_{\alpha i}\left\langle \Delta\mathbf{d}(\Delta\mathbf{d})^T \right\rangle_{ij}(\mathbf{A}^{-g})_{\alpha j}$$

$$= \sum_{\beta=1}\left(\mathbf{A}_{\alpha\beta}^{-g}\sigma_{d_\beta}\right)^2 \qquad (5)$$

由（5）可知，如果广义逆的分量中存在很大的数，则估计模型的误差就可能很大。这是我们不希望看到的。以下讨论广义逆的求法。模型的方差不仅与数据的质量有关还与数据量及数据的结构有关。

讨论：从（4）看，模型的分辨与数据无关吗？

3.5 最小二乘估计

设数据分量的个数大于模型分量的个数。由于数据误差的存在，一般来说线性方程组 $\mathbf{d}=\mathbf{Am}$ 是不成立的，严格的表示应该是 $\mathbf{d}=\mathbf{Am}+\boldsymbol{\varepsilon}$。但在实际观测中我们并不知道准确的误差，所以 $\mathbf{d}=\mathbf{Am}$ 是超定的矛盾方程组，严格的意义下无解。但是我们可以求最小二乘意义下最大限度拟合数据的解 $\tilde{\mathbf{m}}$。先来看一个简单的例子。由两个物体，其质量分别为 m_1 和 m_2。称第一个物体的重量是1千克。称第二个物体的重量是2千克。两个物体一起称得重量为2千克。方程为

$$\begin{cases} m_1 = d_1 = 1 \\ m_2 = d_2 = 2 \\ m_1 + m_2 = d_3 = 2 \end{cases} \tag{6}$$

矩阵 $\mathbf{A}$ 由下式给出

$$\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix} \tag{7}$$

很明显，方程是矛盾的超定系统。因为不可能有第一个物体的重量是1千克，第二个物体的重量2千克，两个物体一起的重量还是2千克。其中一定出现了测量误差。在不知道哪一次测量错误的情况下，我们不可能去除某一个测量数据或去除某一个方程而偏重另两个方程。图3是这一问题的图示。方程组（6）在 (m_1, m_2) 平面上对应着3条直线。三条线不可能相交于一点，表明方程系统不和谐。反演问题归结为以某种方式来协调这一方程组，求出一个最大限度拟合观测数据的解。

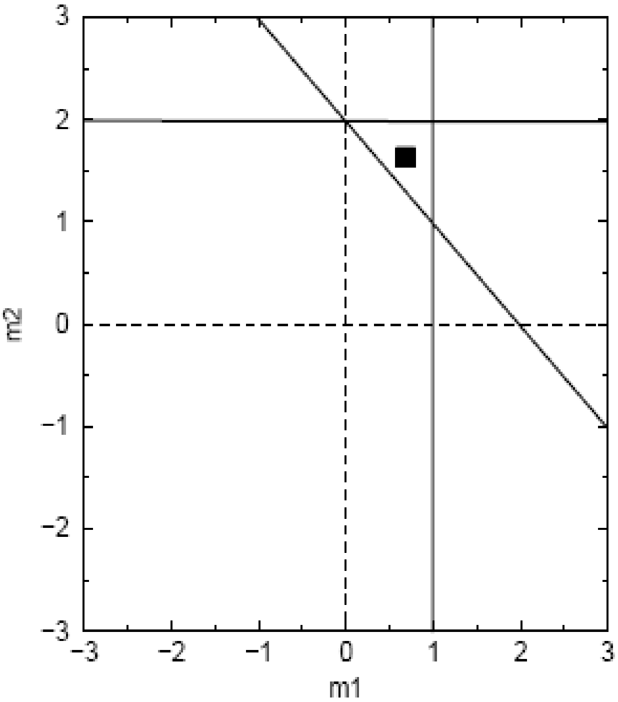

图 3. 线性方程组(6)的几何解释

通常的方法是求出一个模型 $\tilde{\mathbf{m}}$，使得 $\mathbf{d}-\mathbf{A}\tilde{\mathbf{m}}$ 的 L_2 范数最小。即

$$\min\left\|\mathbf{d}-\mathbf{A}\tilde{\mathbf{m}}\right\|^2 \qquad (8)$$

满足条件（8）的解叫最小二乘意义下的解。以下来求线性方程组

$$\mathbf{d}=\mathbf{Am}+\boldsymbol{\varepsilon} \qquad (9)$$

满足条件（8）的模型 $\tilde{\mathbf{m}}$。把（8）写成下角标形式

$$S=(d_i-A_{ij}\tilde{m}_j)(d_i-A_{ik}\tilde{m}_k)$$

求导数

$$\frac{\partial S}{\partial \tilde{m}_l}=-2A_{ij}\delta_{jl}(d_i-A_{ik}\tilde{m}_k)=0$$

推得 $$A_{li}^T d_i=A_{li}^T A_{ik}\tilde{m}_k$$

设 $(\mathbf{A}^T\mathbf{A})^{-1}$存在，则

$$\tilde{\mathbf{m}}=(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{d} \qquad (10)$$

所以方程组（6）的最小二乘解为

$$\tilde{\mathbf{m}} = \frac{1}{3}\begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} \qquad (11)$$

$$\tilde{m}_1 = \frac{2}{3}$$

$$\tilde{m}_2 = \frac{5}{3}$$

图3中的小黑方块就是这一解，他离三条直线的距离最近。

由（10）可知方程组（9）的最小二乘意义下的广义逆为

$$\mathbf{A}^{-g} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$$

根据（4），分辨矩阵

$$\mathbf{R} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{A} = \mathbf{I} \qquad (12)$$

（12）告诉我们，如果$(\mathbf{A}^T\mathbf{A})^{-1}$存在，则最小二乘解为完全分辨。

2．2 最小范数估计

设模型参数的分量个数大于数据分量的个数，并设方程组$\mathbf{d} = \mathbf{A}\mathbf{m}$是和谐的。在此情况下，$(\mathbf{A}^T\mathbf{A})^{-1}$不存在。由于方程数小于未知数个数，方程的解有无穷多个。举个例子，两个物体m_1和m_2的总重量为2，求m_1，m_2的重量。系统方程为

$$m_1 + m_2 = 2 \qquad (13)$$

则

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \end{pmatrix}$$

$$\mathbf{A}^T\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

其逆不存在。图4是方程（13）的解的图示。很明显在直线上的任意点都是方程（13）的解。

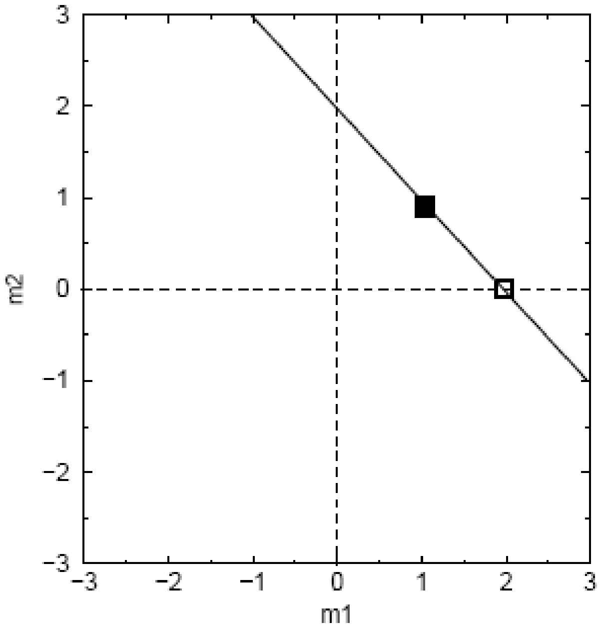

图 4. 方程(13)的解的几何解释

此类问题属于不定方程组解的问题。不定方程组的求解不能用最小二乘准则，因为方程组是和谐的，$\mathbf{d}-\mathbf{A}\tilde{\mathbf{m}}=\mathbf{0}$。一般来说，不定方程组有无穷多组解。我们的准则是求最小 L_2 范数的解 $\tilde{\mathbf{m}}$，即求方程组 $\mathbf{d}=\mathbf{Am}$ 满足条件 $\min \|\mathbf{m}\|^2$ 的解 $\tilde{\mathbf{m}}$。

求 $\|\mathbf{m}\|^2$ 的极小值，满足条件 $\mathbf{d}=\mathbf{Am}$。属于条件极值问题。根据拉格朗日参数法，

令 $$S=\|\mathbf{m}\|^2+\mathbf{\Lambda}^T(\mathbf{d}-\mathbf{Am}) \tag{14}$$

其中 $$\mathbf{\Lambda}^T=(\lambda_1 \quad \lambda_2 \quad \cdots \lambda_n) \tag{15}$$

为拉格朗日乘数矢量。用下角标表示

$$S=m_i^2+\lambda_i(d_i-A_{ij}m_j)$$

求导数并令其为零得

$$\frac{\partial S}{\partial m_k}=2m_k-\lambda_i A_{ik}$$

$$2m_k=\lambda_i A_{ik}$$

即 $$2\mathbf{m}=\mathbf{A}^T\mathbf{\Lambda} \tag{16}$$

进而得 $$2\mathbf{Am}=\mathbf{AA}^T\mathbf{\Lambda}$$

$$2\mathbf{d} = \mathbf{A}\mathbf{A}^T\mathbf{\Lambda} \tag{17}$$

设矩阵 $\mathbf{A}\mathbf{A}^T$ 的逆存在，由（17）得

$$\mathbf{\Lambda} = 2(\mathbf{A}\mathbf{A}^T)^{-1}\mathbf{d} \tag{18}$$

将（18）代入（16）得

$$\tilde{\mathbf{m}} = \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}\mathbf{d} \tag{19}$$

（19）即为模型参数的最小 L_2 范数的解。将其应用于（13）得

$$m_1 = m_2 = 1 \tag{20}$$

矢量的范数

范数的概念是绝对值的推广，一个标量，用 $\|\cdot\|$ 表示。设 $\mathbf{a},\mathbf{b}$ 是相对于数域 F 的线性空间中的任意矢量，矢量的范数具有如下性质

1．$\|\mathbf{a}\| \geq 0$, 仅当 $\mathbf{a} = \mathbf{0}$, 等号才成立

2．$\|\alpha\mathbf{a}\| = |\alpha|\|\mathbf{a}\|$，$\alpha \in F$

3．$\|\mathbf{a}+\mathbf{b}\| \leq \|\mathbf{a}\| + \|\mathbf{b}\|$

n 维实空间 $\mathbf{R}^n$ 中的任意矢量 $\mathbf{a}$ 的 L_p 范数定义为

$$\|\mathbf{a}\|_{L_p} = \left(\sum_{\alpha=1}^{n} |a_\alpha|^p\right)^{\frac{1}{p}}$$

由（19）得广义逆

$$\mathbf{A}^{-g} = \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1} \tag{21}$$

分辨矩阵为

$$\mathbf{R} = \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}\mathbf{A} \tag{22}$$

2．3　混合问题

在运用最小二乘准则时我们假设数据量大于模型参数，方程是超定的矛盾方程组，并且 $(\mathbf{A}^T\mathbf{A})^{-1}$ 存在。在求最小范数解时，我们假定方程组是和谐的，但没有足够的数据量来

确定模型参数，方程组是不定方程组，但$(\mathbf{AA}^T)^{-1}$存在。但是在实际问题中，更一般的情况是，由于误差的存在对于某些模型参数我们有矛盾的数据，而对于另一些模型参数我们又没有足够的数据对其进行估计，并且$(\mathbf{A}^T\mathbf{A})^{-1}$和$(\mathbf{AA}^T)^{-1}$都不存在，问题是病态的。即使矩阵的逆存在，方程组也可能是条件病态的，就是说，数据矢量的微小变化引起模型参数的巨大变化，数据的误差在模型空间中被放大了，解是不稳定的。对于这些情况，之前定义的最小二乘准则和最小范数准则都不能解决。我们必须另辟蹊径解决这些问题。Levenberg（1944）提出了阻尼最小二乘法。从数学的观点考虑，病态的结果是矩阵$\mathbf{A}$的零或接近零的奇异值引起的。

设矩阵$\mathbf{M}$有本征值λ_n和本征向量$\mathbf{v}_n$

$$\mathbf{M}\mathbf{v}_n = \lambda_n \mathbf{v}_n \qquad (23)$$

则矩阵$(\mathbf{M}+\gamma\mathbf{I})$的本征值是$(\lambda_n+\gamma)$

$$(\mathbf{M}+\gamma\mathbf{I})\mathbf{v}_n = (\lambda_n+\gamma)\mathbf{v}_n \qquad (24)$$

这意味着，矩阵的本征值可以通过在其对角线元素上增加一个正的值而加大。这一特性被用来定义阻尼最小二乘解

$$\tilde{\mathbf{m}} = (\mathbf{A}^T\mathbf{A}+\gamma\mathbf{I})^{-1}\mathbf{A}^T\mathbf{d} \qquad (25)$$

由于矩阵$\mathbf{A}^T\mathbf{A}$的本征值大于等零，如果γ的值大于零，则矩阵$(\mathbf{A}^T\mathbf{A})$的零或接近零的本征值的影响就被消除了。

证明$\mathbf{A}^T\mathbf{A}$的本征值≥ 0

证：$\because$ $\mathbf{A}^T\mathbf{A}\mathbf{v}_n = \lambda_n\mathbf{v}_n$，则 $\mathbf{v}_n^T\mathbf{A}^T\mathbf{A}\mathbf{v}_n = \lambda_n\|\mathbf{v}_n\|^2$

$$\lambda_n = \frac{(\mathbf{A}\mathbf{v}_n)^T\mathbf{A}\mathbf{v}_n}{\|\mathbf{v}_n\|^2} = \frac{\|\mathbf{A}\mathbf{v}_n\|^2}{\|\mathbf{v}_n\|^2} \geq 0,$$ 证毕。

由（25）表征的解可由求下式的极小值获得

$$S = \|\mathbf{d}-\mathbf{A}\tilde{\mathbf{m}}\|^2 + \gamma\|\tilde{\mathbf{m}}\|^2 \qquad (26)$$

以下脚标形式表示

$$S = (d_i - A_{ij}\tilde{m}_j)^2 + \gamma\tilde{m}_i^2$$

求导数并令其等于零得

$$\frac{\partial S}{\partial \tilde{m}_k} = 2(d_i - A_{ij}\tilde{m}_j)(-A_{ik}) + 2\gamma\tilde{m}_k = 0$$

即 $$(\mathbf{A}^T\mathbf{A}+\gamma\mathbf{I})\tilde{\mathbf{m}}=\mathbf{A}^T\mathbf{d}$$

所以 $$\tilde{\mathbf{m}}=(\mathbf{A}^T\mathbf{A}+\gamma\mathbf{I})^{-1}\mathbf{A}^T\mathbf{d} \tag{27}$$

由（6）可知，解（27）尽可能的拟合数据同时模型的范数又不会太大。所以γ的作用是一个调和因子。我们来求（27）的另一种表示形式，令

$$\mathbf{Y}=(\mathbf{A}^T\mathbf{A}+\boldsymbol{\gamma}\mathbf{I})^{-1}\mathbf{A}^T \tag{28}$$

推得 $$\mathbf{A}(\mathbf{A}^T\mathbf{A}+\boldsymbol{\gamma}\mathbf{I})\mathbf{Y}=\mathbf{A}\mathbf{A}^T$$

$$(\mathbf{A}\mathbf{A}^T+\boldsymbol{\gamma}\mathbf{I})\mathbf{A}\mathbf{Y}=\mathbf{A}\mathbf{A}^T$$

$$\mathbf{A}\mathbf{Y}=(\mathbf{A}\mathbf{A}^T+\boldsymbol{\gamma}\mathbf{I})^{-1}\mathbf{A}\mathbf{A}^T \tag{29}$$

将（28）带入（29）得

$$\mathbf{A}(\mathbf{A}^T\mathbf{A}+\boldsymbol{\gamma}\mathbf{I})^{-1}\mathbf{A}^T=(\mathbf{A}\mathbf{A}^T+\boldsymbol{\gamma}\mathbf{I})^{-1}\mathbf{A}\mathbf{A}^T$$

所以 $$\mathbf{A}(\mathbf{A}^T\mathbf{A}+\boldsymbol{\gamma}\mathbf{I})^{-1}=(\mathbf{A}\mathbf{A}^T+\boldsymbol{\gamma}\mathbf{I})^{-1}\mathbf{A}$$

转置得 $$(\mathbf{A}^T\mathbf{A}+\boldsymbol{\gamma}\mathbf{I})^{-1}\mathbf{A}^T=\mathbf{A}^T(\mathbf{A}\mathbf{A}^T+\boldsymbol{\gamma}\mathbf{I})^{-1} \tag{30}$$

所以（27）可写为另一种形式

$$\tilde{\mathbf{m}}=\mathbf{A}^T(\mathbf{A}\mathbf{A}^T+\gamma\mathbf{I})^{-1}\mathbf{d} \tag{31}$$

为计算方便，根据不同的问题选择不同的解的形式，尽可能的选择界数小的矩阵求逆。同时我们看到，在阻尼系数不为零的情况下，最小二乘解与最小阻尼解是相同的。

阻尼最小二乘法的广义逆为

$$\mathbf{A}^{-\mathbf{g}}=\mathbf{A}^T(\mathbf{A}\mathbf{A}^T+\gamma\mathbf{I})^{-1} \tag{32}$$

或者 $$\mathbf{A}^{-\mathbf{g}}=(\mathbf{A}^T\mathbf{A}+\gamma\mathbf{I})^{-1}\mathbf{A}^T \tag{33}$$

分辨矩阵为

$$\mathbf{R}=\mathbf{A}^T(\mathbf{A}\mathbf{A}^T+\gamma\mathbf{I})^{-1}\mathbf{A} \tag{34}$$

或者 $$\mathbf{R}=(\mathbf{A}^T\mathbf{A}+\gamma\mathbf{I})^{-1}\mathbf{A}^T\mathbf{A} \tag{35}$$

在讨论最小二乘解的时候我们知道其解是完全分辨的，亦即分辨矩阵为单位矩阵。但由（35）可知，只要阻尼因子不为零，阻尼最小二乘的分辨矩阵不可能是单位矩阵，也就不可能完全分辨。这意味着阻尼最小二乘解降低了分辨。阻尼因子γ越大，分辨就越低。另一方面，由于矩阵$\mathbf{A}$的奇异值（零或接近零的值）只出现在广义逆元素的分母里，当γ增大时分母就越大，从而减小了模型参数的误差。所以，γ不仅在最小二乘与最小范数之间起调和作用，真正的意义是在模型参数的分辨与误差之间调和。

思考：当γ=0 时，会发生什么情况。

2．4　 最小二乘解中的协调性问题

对于矛盾超定的方程组，最小二乘方法是乎提供了一个客观公正的解。然而，实际上还是存在一些问题。我们先来看一个例子。方程组（6）的第三式的两边都乘上2得，

$$\begin{cases} m_1 = d_1 = 1 \\ m_2 = d_2 = 2 \\ m_1 + m_2 = d_3 = 2 \end{cases} \Leftrightarrow \begin{cases} m_1 = d_1 = d_1' = 1 \\ m_2 = d_2 = d_2' = 2 \\ 2m_1 + 2m_2 = 2d_3 = d_3' = 4 \end{cases} \qquad (36)$$

数学上这两个系统是完全等价的。他们的系统矩阵为

$$\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}, \qquad \mathbf{A}' = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 2 & 2 \end{pmatrix} \qquad (37)$$

带撇的量是变换后的系统。利用（10）很容易求出两个系统的最小二乘解

$$\tilde{\mathbf{m}} = \frac{1}{3}\begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \end{pmatrix}\mathbf{d}, \qquad \tilde{\mathbf{m}}' = \frac{1}{9}\begin{pmatrix} 5 & -4 & 2 \\ -4 & 5 & 2 \end{pmatrix}\mathbf{d}' \qquad (38)$$

$$\tilde{\mathbf{m}} = \begin{pmatrix} 2/3 \\ 5/3 \end{pmatrix}, \qquad \tilde{\mathbf{m}}' = \begin{pmatrix} 5/9 \\ 14/9 \end{pmatrix} \qquad (39)$$

问题出现了，数学上两个等价的系统其最小二乘解怎么会不相同呢！任意的系统变换都会出现不同的最小二乘解。所以，最小二乘解并不像表面上看起来那样客观公正。那么问题到底出现在哪里呢。

最小二乘准则是解矛盾的超定方程组。为什么会出现矛盾的方程呢。其根本原因是我们并不知道真正的误差到底是多少。和谐的方程应该是

$$\mathbf{d} = \mathbf{Am} + \boldsymbol{\varepsilon} \qquad (40)$$

由于我们不知道误差矢量$\boldsymbol{\varepsilon}$，于是我们就简单的求解方程

$$\mathbf{d} = \mathbf{Am} \qquad (41)$$

。去掉误差的方程（41）是矛盾方程！问题就出现在误差传播项$\mathbf{A}^{-g}\boldsymbol{\varepsilon}$上了。还是以系统（6）为例。假设物体$m_1$的称重为1千克是准确的，$m_2$为两千克也是准确的。那么误差就出现在第三次称重上。和谐的方程应该是

$$\begin{cases} m_1 = d_1 = 1 \\ m_2 = d_2 = 2 \\ m_1 + m_2 = d_3 = 2 + \varepsilon_3 \end{cases} \qquad (42)$$

其中$\varepsilon_3 = 1$。如果（42）的第三式两边乘上2，那么误差项也乘上了两倍，方程就还是和谐的。但误差被放大了两倍。但在求新的系统最小二乘解的时候，没有考虑到误差的放大。从而使得数据空间中的标度发生了变化。所以，变换后的系统与原来的系统并非真正等价！由于误差的存在而产生了矛盾方程系统，最终导致最小二乘解对标度操作的依赖性。

下面我们在一般意义上进一步详细讨论变换后的系统最小二乘解的性质。最初的系统为

$$\mathbf{d} = \mathbf{Am} \qquad (43)$$

我们记住，由于我们去除了误差项，所以（43）并不严格成立，它是矛盾的超定系统。我们对模型参数进行变换

$$\mathbf{m}' = \mathbf{Sm} \qquad （44）$$

对数据进行变换

$$\mathbf{d}' = \mathbf{Qd} \qquad （45）$$

设变换矩阵$\mathbf{S}$的逆存在。（43）通过变换后为

$$\mathbf{QAS}^{-1}\mathbf{m}' = \mathbf{Qd} = \mathbf{d}' \qquad （46）$$

系统（43）的最小二乘解为

$$\tilde{\mathbf{m}}^{(1)} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{d} \qquad （47）$$

令 $$\mathbf{A}' = \mathbf{QAS}^{-1} \qquad （48）$$

方程组（46）变为

$$\mathbf{d}' = \mathbf{A}'\mathbf{m}' \qquad （49）$$

其最小二乘解为

$$\tilde{\mathbf{m}}' = (\mathbf{A}'^T\mathbf{A}')^{-1}\mathbf{A}'^T\mathbf{d}' \qquad （50）$$

将（44），（45），（48）代入（50）

$$\tilde{\mathbf{m}}^{(2)} = \mathbf{S}^{-1}(\mathbf{S}^{-T}\mathbf{A}^T\mathbf{Q}^T\mathbf{QAS}^{-1})^{-1}\mathbf{S}^{-T}\mathbf{A}^T\mathbf{Q}^T\mathbf{Qd} \qquad （51）$$

令 $$\mathbf{Y} = \mathbf{S}^{-1}(\mathbf{S}^{-T}\mathbf{A}^T\mathbf{Q}^T\mathbf{QAS}^{-1})^{-1}\mathbf{S}^{-T} \qquad （52）$$

则 $$\mathbf{S}^T(\mathbf{S}^{-T}\mathbf{A}^T\mathbf{Q}^T\mathbf{QAS}^{-1})\mathbf{SY} = \mathbf{I}$$

所以 $$\mathbf{Y} = (\mathbf{A}^T\mathbf{Q}^T\mathbf{QAI})^{-1} \qquad （53）$$

利用（52）（53），（51）简化为

$$\tilde{\mathbf{m}}^{(2)} = (\mathbf{A}^T\mathbf{Q}^T\mathbf{QA})^{-1}\mathbf{A}^T\mathbf{Q}^T\mathbf{Qd} \qquad （54）$$

比较两个解（54）和（47），只要$\mathbf{Q}^T\mathbf{Q} \neq \mathbf{I}$，这两个解就不相等。但从（54）可知，模型的变换不会影响最小二乘解的结果，只要模型空间的变换是可逆的。这很好理解，因为最小二乘法则是最大限度的拟合数据，而不要求模型的范数最小。影响结果的是数据空间的变换，除非数据空间的变换是幺正变换。从线性代数我们知道，幺正变换是使矢量旋转，矢量的长度不变。也就是变换后空间的标度没变。以上分析表明，对于矛盾的超定系统，数据空间的变换会影响空间标度。导致变换前的结果与变换后的结果不一致。

2．5　最小范数解中的协调性问题

像最小二乘解一样，最小范数解同样存在协调性问题。以不定方程（13）解为例，最小

范数解为

$$\tilde{m}_1 = 1, \quad \tilde{m}_2 = 1 \tag{55}$$

作模型空间的变换

$$\begin{aligned} m_1' &= m_1 + m_2 \\ m_2' &= m_2 \end{aligned} \tag{56}$$

对于新的模型矢量，方程系统变为

$$m_1' = 2 \tag{57}$$

变换后的方程系统的解更加不定，因为系统对参数m_2'没有任何约束。（57）的最小范数解为

$$\tilde{m}_1' = 2, \quad \tilde{m}_2' = 0 \tag{58}$$

这一解在图4中表为空心正方形。很明显，变换前与变换后的解是不相同的。其原因与最小二乘解中的相似，变换（56）改变了模型空间中的标度。两个空间对距离的度量是不相同的。所以它们的解是不相同的。我们可以像在最小二乘解的讨论时一样，作一般的变换分析。由于阻尼最小二乘解是最小二乘与最小范数之间的协调，在阻尼系数不为零的情况下，最小二乘解与最小范数解是相同的。所以，更一般的讨论应该放到阻尼最小二乘中去做。

2．6　阻尼最小二乘解中协调性讨论

忽略误差，方程系统为$\mathbf{d} = \mathbf{A}\mathbf{m}$，令$\mathbf{d}' = \mathbf{Q}\mathbf{d}, \quad \mathbf{m}' = \mathbf{S}\mathbf{m}$，并且设$\mathbf{S}^{-1}$存在。变换后的系统为 $\mathbf{Q}\mathbf{A}\mathbf{S}^{-1}\mathbf{m}' = \mathbf{Q}\mathbf{d}$。变换前的阻尼最小二乘解为

$$\tilde{\mathbf{m}}^{(1)} = (\mathbf{A}^T\mathbf{A} + \gamma\mathbf{I})^{-1}\mathbf{A}^T\mathbf{d} \tag{59}$$

变换后的阻尼最小二乘解为

$$\tilde{\mathbf{m}}^{(2)} = (\mathbf{A}^T\mathbf{Q}^T\mathbf{Q}\mathbf{A} + \gamma\mathbf{S}^T\mathbf{S})^{-1}\mathbf{A}^T\mathbf{Q}^T\mathbf{Q}\mathbf{d} \tag{60}$$

比较（59）与（60）可知，除非数据空间和模型空间都是幺正变换，否则它们不相等。从(60)还可看出，在$\mathbf{m}'$空间中，阻尼项是$\gamma\mathbf{I}$，变换回到$\mathbf{m}$空间后，阻尼项变成$\gamma\mathbf{S}^T\mathbf{S}$。一般来说，阻尼最小二乘解在数据空间和模型空间变换后不是不变量。

在实际处理资料时我们并不能判断变换前后的两个解哪一个更合理，它们在数学上是平等的。这个问题非常严峻。变换前后的解必须一致是反演的基本要求。否则就有无穷多个不同的解，而且它们都是平等的。于是，我们必须寻求一种合理的法则，使得变换后其解不变。设$\mathbf{W}_d = \mathbf{Q}^T\mathbf{Q}$，$\mathbf{W}_m = \mathbf{S}^T\mathbf{S}$，（60）变为

$$\tilde{\mathbf{m}} = (\mathbf{A}^T\mathbf{W}_d\mathbf{A} + \gamma\mathbf{W}_m)^{-1}\mathbf{A}^T\mathbf{W}_d\mathbf{d} \tag{61}$$

解（61）是最小二乘解更一般的表达形式，称为广义最小二乘解。它可以通过求解下面的价值函数的极小获得

$$S = (\mathbf{d} - \mathbf{Am})^T \mathbf{W}_d (\mathbf{d} - \mathbf{Am}) + \gamma \mathbf{m}^T \mathbf{W}_m \mathbf{m} \qquad (62)$$

其中 $\mathbf{W}_d$ 和 $\mathbf{W}_m$ 都是对称矩阵。这是最小二乘准则更一般的表达形式。

我们讨论在模型空间和数据空间作变换后 $\mathbf{W}_d$ 和 $\mathbf{W}_m$ 作什么样的变换才能使得变换后的解不变。设数据空间和模型空间的变换为 $\mathbf{d}' = \mathbf{Qd},\quad \mathbf{m}' = \mathbf{Sm}$，并设 $\mathbf{Q}^{-1}$和$\mathbf{S}^{-1}$存在，则 $\mathbf{A}' = \mathbf{QAS}^{-1}$，代入（62）得

$$S = (\mathbf{d}' - \mathbf{A}'\mathbf{m}')^T \mathbf{Q}^{-T} \mathbf{W}_d \mathbf{Q}^{-1} (\mathbf{d}' - \mathbf{A}'\mathbf{m}') + \gamma \mathbf{m}'^T \mathbf{S}^{-T} \mathbf{W}_m \mathbf{S}^{-1} \mathbf{S} \mathbf{m}' \qquad (63)$$

令
$$\begin{aligned} \mathbf{W}_d' &= \mathbf{Q}^{-T} \mathbf{W}_d \mathbf{Q}^{-1} \\ \mathbf{W}_m' &= \mathbf{S}^{-T} \mathbf{W}_m \mathbf{S}^{-1} \end{aligned} \qquad (64)$$

带入（63）得

$$S = (\mathbf{d}' - \mathbf{A}'\mathbf{m}')^T \mathbf{W}_d' (\mathbf{d}' - \mathbf{A}'\mathbf{m}') + \gamma \mathbf{m}'^T \mathbf{W}_m' \mathbf{m}' \qquad (65)$$

由此可知，$\mathbf{W}_d$ 和 $\mathbf{W}_m$ 只要按照（64）进行变换，价值函数的形式和值都不变。变换后 $\mathbf{m}'$ 的广义阻尼最小二乘解为

$$\tilde{\mathbf{m}}' = (\mathbf{A}'^T \mathbf{W}_d' \mathbf{A}' + \gamma \mathbf{W}_m')^{-1} \mathbf{A}'^T \mathbf{W}_d' \mathbf{d}' \qquad (66)$$

容易验证，变回到 $\mathbf{m}$和$\mathbf{d}$ 空间后，（66）就变为（61）。这就满足了变换前后的解相同的要求。阻尼因子 γ 可以归并到权矩阵 $\mathbf{W}_m$ 中，现在总结一下，变换前后的解相同的条件，

（i）坐标变换 $\mathbf{d}' = \mathbf{Qd},\quad \mathbf{m}' = \mathbf{Sm}$，设 $\mathbf{Q}^{-1}$和$\mathbf{S}^{-1}$存在，新的系统矩阵 $\mathbf{A}' = \mathbf{QAS}^{-1}$

(ii) 加权矩阵的变换 $\begin{aligned} \mathbf{W}_d' &= \mathbf{Q}^{-T} \mathbf{W}_d \mathbf{Q}^{-1} \\ \mathbf{W}_m' &= \mathbf{S}^{-T} \mathbf{W}_m \mathbf{S}^{-1} \end{aligned}$

(iii) 价值函数 $S = (\mathbf{d} - \mathbf{Am})^T \mathbf{W}_d (\mathbf{d} - \mathbf{Am}) + \mathbf{m}^T \mathbf{W}_m \mathbf{m}$

变换后的价值函数 $S = (\mathbf{d}' - \mathbf{A}'\mathbf{m}')^T \mathbf{W}_d' (\mathbf{d}' - \mathbf{A}'\mathbf{m}') + \mathbf{m}'^T \mathbf{W}_m' \mathbf{m}'$

(iv) 求价值函数的极小值

现在再回到2.4中举的例子，看看如何具体解决变换前后解不一致的问题。当时的解释是，数据空间变换后，误差随之而变，由于没有进行协调，从而造成了变换前后数据空间的标度不一致，最终导致解的不一致。在那个例子中， $\mathbf{W}_m = \mathbf{0}$， $\mathbf{W}_d = \mathbf{I}$，数据空间的变换矩阵为

$$\mathbf{Q}=\begin{pmatrix}1&0&0\\0&1&0\\0&0&2\end{pmatrix},\quad \mathbf{Q}^{-1}=\begin{pmatrix}1&0&0\\0&1&0\\0&0&\frac{1}{2}\end{pmatrix} \qquad (67)$$

变换前的价值函数为

$$S=(\mathbf{d}-\mathbf{Am})^T(\mathbf{d}-\mathbf{Am}) \qquad (68)$$

根据条件（ii）和（iii），变换后的价值函数应为

$$S=(\mathbf{d}'-\mathbf{A}'\mathbf{m}')^T\mathbf{Q}^{-T}\mathbf{Q}^{-1}(\mathbf{d}'-\mathbf{A}'\mathbf{m}') \qquad (69)$$

而不是 $S'=(\mathbf{d}'-\mathbf{A}'\mathbf{m}')^T(\mathbf{d}'-\mathbf{A}'\mathbf{m}')$

这就是变换前后解不一致的根本所在。用误差矢量表示（69）得

$$S=\boldsymbol{\varepsilon}'^T\mathbf{Q}^{-T}\mathbf{Q}^{-T}\boldsymbol{\varepsilon}' \qquad (70)$$

根据方程（42），变换后的误差 $\varepsilon_1'=\varepsilon_1,\varepsilon_2'=\varepsilon_2$，$\varepsilon_3'=2\varepsilon_3$，误差被放大了两倍，。再看（70），把（67）代入（70）

$$S=\begin{pmatrix}\varepsilon_1 & \varepsilon_2 & \frac{1}{2}\varepsilon_3'\end{pmatrix}\begin{pmatrix}\varepsilon_1\\ \varepsilon_2\\ \frac{1}{2}\varepsilon_3'\end{pmatrix}=\begin{pmatrix}\varepsilon_1 & \varepsilon_2 & \varepsilon_3\end{pmatrix}\begin{pmatrix}\varepsilon_1\\ \varepsilon_2\\ \varepsilon_3\end{pmatrix} \qquad (71)$$

上式表明，$\mathbf{Q}^{-T}\mathbf{Q}^{-1}$ 把放大的误差调整回去了。从而没有改变变换后的空间度量。

上述分析告诉我们，在实际工作中应尽量采用原始测量数据来建立方程系统。如果有必要做数据和模型的变换，那么权重矩阵 $\mathbf{W}_d$ 和 $\mathbf{W}_m$ 要按照条件(ii)作相应的变换。否则将会改变数据误差，导致数据空间标度的改变。

具体工作中如何选择 $\mathbf{W}_d$ 和 $\mathbf{W}_m$，没有固定的法则，并且带有很大的主观性。一般来说，它们应该包含模型和数据的统计先验信息，对数据和模型参数进行加权。以后在讲到弱非线性问题的反演时，我们再讨论它们的具体选择。

2．7 矩阵的奇异值分解

求解线性方程系统 $\mathbf{d}=\mathbf{Am}$ 的关键是求广义逆 $\mathbf{A}^{-g}$。有了广义逆，不仅能够得到估计模型，而且能够评估模型。利用最小二乘准则，我们已经知道广义逆的求法。但在解的非唯一性以及数据的误差分析上没有很清晰的图象。通过矩阵的奇异值分解，能够帮助我们解决这些问题。

设矩阵 $\mathbf{A}$ 为 $n\times m$ 阶矩阵。根据Lanczos（1961）的方法，构造一个对称矩阵 $\mathbf{S}$。它由矩阵 $\mathbf{A}$ 和 $\mathbf{A}^T$ 构造而成，

$$\mathbf{S}=\begin{pmatrix}\mathbf{0} & \mathbf{A}\\ \mathbf{A}^T & \mathbf{0}\end{pmatrix} \qquad (2.72)$$

矩阵$\mathbf{S}$是$(n+m)\times(n+m)$阶对称方阵。根据线性代数，$\mathbf{S}$有$n\times m$个相互正交的本征矢量$\mathbf{w}_\alpha\ (\alpha=1,2,\cdots,n+m)$，本征值$\lambda_\alpha$是实数，满足

$$\mathbf{S}\mathbf{w}_\alpha=\lambda_\alpha\mathbf{w}_\alpha \qquad (i=1,2,\cdots,n+m) \qquad (2.73)$$

由于本征矢量$\mathbf{w}_\alpha$ 有$n+m$个分量，我们把它分成两部分，$\mathbf{u}_\alpha$和$\mathbf{v}_\alpha$，即

$$\mathbf{w}_\alpha=\begin{pmatrix}\mathbf{u}_\alpha\\ \mathbf{v}_\alpha\end{pmatrix} \qquad (2.74)$$

其中$\mathbf{u}_\alpha$是数据空间中的n维矢量，$\mathbf{v}_\alpha$是模型空间中的m维矢量。如此分解后，（2.73）可表为

$$\begin{pmatrix}\mathbf{0} & \mathbf{A}\\ \mathbf{A}^T & \mathbf{0}\end{pmatrix}\begin{pmatrix}\mathbf{u}_\alpha\\ \mathbf{v}_\alpha\end{pmatrix}=\lambda_\alpha\begin{pmatrix}\mathbf{u}_\alpha\\ \mathbf{v}_\alpha\end{pmatrix} \qquad (2.75)$$

对于非零本征值λ_α，由（2.75）得到关于矢量$\mathbf{u}_\alpha$和$\mathbf{v}_\alpha$的耦合方程

$$\begin{cases}\mathbf{A}\mathbf{v}_\alpha=\lambda_\alpha\mathbf{u}_\alpha\\ \mathbf{A}^T\mathbf{u}_\alpha=\lambda_\alpha\mathbf{v}_\alpha\end{cases}, \quad (\alpha=1,2\cdots p) \qquad (2.76)$$

对于零本征值，$\mathbf{u}_\alpha$和$\mathbf{v}_\alpha$之间解耦，设非零本征值的个数为p，则

$$\begin{aligned}\mathbf{A}\mathbf{v}_\alpha&=\mathbf{0}, \quad (\alpha=p+1,\cdots m)\\ \mathbf{A}^T\mathbf{u}_\alpha&=\mathbf{0}, \quad (\alpha=p+1,\cdots m)\end{aligned} \qquad (2.77)$$

由(2.76)进一步得到

$$\begin{aligned}\mathbf{A}^T\mathbf{A}\mathbf{v}_\alpha&=\lambda_\alpha^2\mathbf{v}_\alpha\\ \mathbf{A}\mathbf{A}^T\mathbf{u}_\alpha&=\lambda_\alpha^2\mathbf{u}_\alpha\end{aligned} \qquad (2.78)$$

由于$\mathbf{A}^T\mathbf{A}$是对称矩阵，所以存在m个相互正交的归一化本征矢量组

$$\mathbf{v}_\alpha\ (\alpha=1,2,\cdots m),$$

并且其本征值是大于等于零的实数。$\mathbf{v}_\alpha\ (\alpha=1,2,\cdots m)$在模型空间中构成一完备正交基

$$\mathbf{v}_\alpha^T\mathbf{v}_\beta=\delta_{\alpha\beta}, (\alpha,\beta=1,2,\cdots,m) \qquad (2.79)$$

其中有p个单位矢量对应于非零本征值λ_α（$\alpha=1,2,\cdots$，p），$m-p$个单位矢量对应于零

本征值。这样，模型空间就被分为两个子空间。一个由非零本征值的本征单位矢量组 $\mathbf{v}_\alpha\ (\alpha=1,2,\cdots p)$ 生成，另一个由零本征值的本征单位矢量组 $\mathbf{v}_\alpha\ (\alpha=p+1,\cdots m)$ 生成。并且这两个子空间是相互正交的。

同样，$\mathbf{AA}^T$ 是对称矩阵，所以存在 n 个相互正交的归一化本征矢量组

$$\mathbf{u}_\alpha\ (\alpha=1,2,\cdots n),$$

并且其本征值是大于等于零的实数。$\mathbf{u}_\alpha\ (\alpha=1,2,\cdots n)$ 在数据空间中构成一完备正交基

$$\mathbf{u}_\alpha^T\mathbf{u}_\beta=\delta_{\alpha\beta}\ ,(\alpha,\beta=1,2,\cdots,n) \qquad (2.80)$$

其中有 p 个单位矢量对应于非零本征值 λ_α（α=1, 2,⋯，p），$n-p$ 个单位矢量对应于零本征值。这样，数据空间也被分为两个子空间。一个由非零本征值的本征单位矢量组 $\mathbf{u}_\alpha\ (\alpha=1,2,\cdots p)$ 生成，另一个由零本征值的本征单位矢量组 $\mathbf{u}_\alpha\ (\alpha=p+1,\cdots n)$ 生成。并且这两个子空间是相互正交的。

以 $\mathbf{v}_\alpha\ (\alpha=1,2,\cdots m)$ 为列向量构成矩阵

$$\mathbf{V}=(\mathbf{v}_1\cdots\mathbf{v}_\alpha\cdots\mathbf{v}_m) \qquad (2.81)$$

以非零本征值 λ_α（α=1, 2,⋯，p）的本征矢量 $\mathbf{v}_\alpha\ (\alpha=1,2,\cdots p)$ 为列向量构成矩阵

$$\mathbf{V}_p=(\mathbf{v}_1\cdots\mathbf{v}_\alpha\cdots\mathbf{v}_p) \qquad (2.82)$$

以零本征的矢量 $\mathbf{v}_\alpha\ (\alpha=p+1,\cdots m)$ 为列向量构成矩阵

$$\mathbf{V}_0=(\mathbf{v}_{p+1}\cdots\mathbf{v}_\alpha\cdots\mathbf{v}_m) \qquad (2.83)$$

则
$$\mathbf{V}=\begin{pmatrix}\mathbf{V}_p & \mathbf{V}_0\end{pmatrix} \qquad (2.84)$$

同样，以 $\mathbf{u}_\alpha\ (\alpha=1,2,\cdots n)$ 为列向量构成矩阵

$$\mathbf{U}=(\mathbf{u}_1\cdots\mathbf{u}_\alpha\cdots\mathbf{u}_n) \qquad (2.85)$$

以非零本征值 λ_α（α=1, 2,⋯，p）的本征矢量 $\mathbf{u}_\alpha\ (\alpha=1,2,\cdots p)$ 为列向量构成矩阵

$$\mathbf{U}_p=(\mathbf{u}_1\cdots\mathbf{u}_\alpha\cdots\mathbf{u}_p) \qquad (2.86)$$

以零本征值的矢量 $\mathbf{u}_\alpha\ (\alpha=p+1,\cdots n)$ 为列向量构成矩阵

$$\mathbf{U}_0=(\mathbf{u}_{p+1}\cdots\mathbf{u}_\alpha\cdots\mathbf{u}_m) \qquad (2.87)$$

则 $$\mathbf{U}=\begin{pmatrix}\mathbf{U}_p & \mathbf{U}_0\end{pmatrix} \tag{2.88}$$

根据（2.79）(2.80)得

$$\begin{aligned}\mathbf{U}^T\mathbf{U}=\mathbf{U}\mathbf{U}^T=\mathbf{I}\\ \mathbf{V}^T\mathbf{V}=\mathbf{V}\mathbf{V}^T=\mathbf{I}\end{aligned} \qquad (2,89)$$

$$\begin{aligned}\mathbf{U}_p^T\mathbf{U}_p=\mathbf{I}\\ \mathbf{V}_p^T\mathbf{V}_p=\mathbf{I}\end{aligned} \qquad （2.90）$$

但是 $$\begin{aligned}\mathbf{U}_p\mathbf{U}_p^T\neq\mathbf{I}\\ \mathbf{V}_p\mathbf{V}_p^T\neq\mathbf{I}\end{aligned} \qquad （2.91）$$

由（2.76）（2.77）得

$$\mathbf{AV}=\mathbf{A}\begin{pmatrix}\mathbf{V}_p & \mathbf{V}_0\end{pmatrix}=\begin{pmatrix}\mathbf{AV}_p & \mathbf{AV}_0\end{pmatrix}=\begin{pmatrix}\mathbf{U}_p\mathbf{\Lambda}_p & \mathbf{0}\end{pmatrix}$$

$$=\begin{pmatrix}\mathbf{U}_p & \mathbf{U}_0\end{pmatrix}\begin{pmatrix}\mathbf{\Lambda}_p & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix} \tag{2.92}$$

其中$\mathbf{\Lambda}_p$是以λ_α（α=1, 2,⋯，p）为对角元素的对角矩阵。根据（2.89），用$\mathbf{V}^T$右乘（2.92）得

$$\mathbf{A}=\begin{pmatrix}\mathbf{U}_p & \mathbf{U}_0\end{pmatrix}\begin{pmatrix}\mathbf{\Lambda}_p & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\begin{pmatrix}\mathbf{V}_p^T\\ \mathbf{V}_0^T\end{pmatrix} \tag{2.93}$$

最后得

$$\mathbf{A}=\mathbf{U}\mathbf{\Lambda}\mathbf{V}^T=\mathbf{U}_p\mathbf{\Lambda}_p\mathbf{V}_p^T \tag{2.94}$$

其中 $$\mathbf{\Lambda}=\begin{pmatrix}\mathbf{\Lambda}_p & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix} \qquad （2.95）$$

表达式（2.94）即为矩阵$\mathbf{A}$的奇异值分解（SVD），λ_α（α=1, 2,⋯，p）被称为矩阵$\mathbf{A}$的奇异值。

2.8 SVD广义逆

在介绍SVD之前，我们先讨论矩阵奇异值分解的意义。奇异值分解后，数据空间$\mathbf{U}$被分为两个子空间$\mathbf{U}_p$和$\mathbf{U}_0$，模型空间$\mathbf{V}$也被分为两个子空间$\mathbf{V}_p$和$\mathbf{V}_0$。

由于本征矢量$\mathbf{u}_\alpha$ $(\alpha=1,2,\cdots n)$在数据空间中构成一完备正交基,数据空间中的任意矢量$\mathbf{a}$都可以用这组基展开

$$\mathbf{a}=(\alpha_1\mathbf{u}_1+\cdots+\alpha_p\mathbf{u}_p)+(\alpha_{p+1}\mathbf{u}_{p+1}+\cdots+\alpha_n\mathbf{u}_n)=\mathbf{U}_p\boldsymbol{\alpha}_p+\mathbf{U}_0\boldsymbol{\alpha}_0 \qquad (2.96)$$

$$\boldsymbol{\alpha}_p = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_p \end{pmatrix} \quad , \quad \boldsymbol{\alpha}_0 = \begin{pmatrix} \alpha_{p+1} \\ \vdots \\ \alpha_n \end{pmatrix} \qquad (2.97)$$

$\boldsymbol{\alpha}_p$ 是 $\mathbf{a}$ 在 $\mathbf{U}_p$ 空间中的坐标，$\boldsymbol{\alpha}_0$ 是 $\mathbf{a}$ 在 $\mathbf{U}_0$ 空间中的坐标

同样 $\mathbf{v}_\alpha\ (\alpha = 1, 2, \cdots m)$ 在模型空间中构成一完备正交基, 模型空间中的任意矢量 $\mathbf{b}$ 可以用这组基展开

$$\mathbf{b} = (\beta_1 \mathbf{v}_1 + \cdots + \beta_p \mathbf{v}_p) + (\beta_{p+1} \mathbf{v}_{p+1} + \cdots + \beta_m \mathbf{v}_m) = \mathbf{V}_p \boldsymbol{\beta}_p + \mathbf{V}_0 \boldsymbol{\beta}_0 \qquad (2.98)$$

$$\boldsymbol{\beta}_p = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_p \end{pmatrix} \quad , \quad \boldsymbol{\beta}_0 = \begin{pmatrix} \beta_{p+1} \\ \vdots \\ \beta_m \end{pmatrix} \qquad (2.99)$$

$\boldsymbol{\beta}_p$ 是 $\mathbf{b}$ 在 $\mathbf{V}_p$ 空间中的坐标，$\boldsymbol{\beta}_0$ 是 $\mathbf{b}$ 在 $\mathbf{V}_0$ 空间中的坐标

由（2.76）（2.77）

$$\mathbf{A}\mathbf{b} = \mathbf{A}\mathbf{V}_p \boldsymbol{\beta}_p + \mathbf{A}\mathbf{V}_0 \boldsymbol{\beta}_0 = \mathbf{U}_p \boldsymbol{\Lambda}_p \boldsymbol{\beta}_p$$

$$= \beta_1 \lambda_1 \mathbf{u}_1 + \beta_2 \lambda_2 \mathbf{u}_2 + \cdots + \beta_p \lambda_p \mathbf{u}_p \qquad (2.100)$$

(2.100)表明，算子 $\mathbf{A}$ 把模型空间的任意矢量只映射到数据子空间 $\mathbf{U}_p$，而没有 $\mathbf{U}_0$ 空间分量。又由于 $\mathbf{A}\mathbf{V}_0 = \mathbf{0}$, 实际上算子 $\mathbf{A}$ 只把模型子空间 $\mathbf{V}_p$ 中的矢量映射到数据子空间 $\mathbf{U}_p$，而且映象和像源之间是一一对应的。它们之间是广义的“满射”。

$$\boldsymbol{\beta}_p = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_p \end{pmatrix} \underset{\mathbf{A}^{-g}}{\overset{\mathbf{A}}{\rightleftarrows}} \boldsymbol{\alpha}_p = \begin{pmatrix} \beta_1 \lambda_1 \\ \vdots \\ \beta_p \lambda_p \end{pmatrix} \qquad (2.101)$$

$\mathbf{V}_0$，$\mathbf{U}_0$ 是算子 $\mathbf{A}$ 的盲点，它们之间没有任何联系。如果 $\mathbf{m}$ 是线性方程系统 $\mathbf{A}\mathbf{m} = \mathbf{d}$ 的解，那么 $\mathbf{m} + \mathbf{V}_0 \boldsymbol{\beta}_0$ 也是它的解，其中 $\mathbf{V}_0 \boldsymbol{\beta}_0$ 是子空间 $\mathbf{V}_0$ 中的任意矢量。所以，模型子空间 $\mathbf{V}_0$ 的存在是线性方程系统解不唯一的根源。另一方面，如果数据 $\mathbf{d}$ 中含有 $\mathbf{U}_0$ 分量，那么任何模型矢量 $\mathbf{m}$ 都不可能满足线性方程系统，因为算子 $\mathbf{A}$ 只把模型矢量 $\mathbf{m}$ 映射到 $\mathbf{U}_p$ 空间。所以 $\mathbf{U}_0$ 子空间的存在是线性方程系统矛盾的根源。如果数据 $\mathbf{d}$ 中含有 $\mathbf{U}_0$ 分量，方程组 $\mathbf{A}\mathbf{m} = \mathbf{d}$ 就是矛盾线性方程系统，不可能有严格的解。图5是模型空间在算子的作用下与数据空间之间的关系。

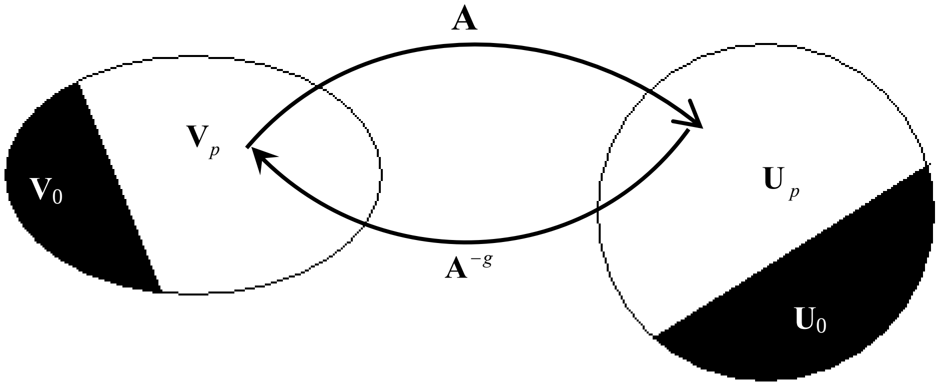

图5 模型空间在算子$\mathbf{A}$的作用下与数据空间的关系，
黑色区是不被算子点亮的盲点

既然模型空间$\mathbf{V}_p$通过算子$\mathbf{A}$映射到数据空间$\mathbf{U}_p$是“满射”，那么就一定存在广义逆算子$\mathbf{A}^{-g}$把数据空间$\mathbf{U}_p$映射到模型空间$\mathbf{V}_p$。我们来求这样的逆算子。设数据$\mathbf{d}$的展开为

$$\mathbf{d} = \begin{pmatrix}\mathbf{u}_1 & \cdots & \mathbf{u}_p\end{pmatrix}\begin{pmatrix}\alpha_1 \\ \vdots \\ \alpha_p\end{pmatrix} + \begin{pmatrix}\mathbf{u}_{p+1} & \cdots & \mathbf{u}_n\end{pmatrix}\begin{pmatrix}\alpha_{p+1} \\ \vdots \\ \alpha_n\end{pmatrix} = \mathbf{U}_p\boldsymbol{\alpha}_p + \mathbf{U}_0\boldsymbol{\alpha}_0 \tag{2.102}$$

则
$$\mathbf{d} - \mathbf{U}_0\boldsymbol{\alpha}_0 = \mathbf{U}_p\boldsymbol{\alpha}_p \tag{2.103}$$

根据（2.101），数据矢量$\mathbf{U}_p\boldsymbol{\alpha}_p$逆射到模型空间$\mathbf{V}_p$中的像是

$$\tilde{\mathbf{m}}_p = \begin{pmatrix}\mathbf{v}_1 & \cdots & \mathbf{v}_p\end{pmatrix}\begin{pmatrix}\alpha_1/\lambda_1 \\ \vdots \\ \alpha_p/\lambda_p\end{pmatrix} = \mathbf{V}_p\boldsymbol{\Lambda}^{-1}\boldsymbol{\alpha}_p \tag{2.104}$$

由于$\mathbf{U}_p^T\mathbf{U}_p = \mathbf{I}$,利用（2.103）得

$$\boldsymbol{\alpha}_p = \mathbf{U}_p^T(\mathbf{d} - \mathbf{U}_0\boldsymbol{\alpha}_0) \tag{2.105}$$

将（2.105）代入（2.104）得

$$\tilde{\mathbf{m}}_p = \mathbf{V}_p\boldsymbol{\Lambda}_p^{-1}\mathbf{U}_p^T(\mathbf{d} - \mathbf{U}_0\mathbf{d}_0)$$

由于 $\mathbf{U}_p \perp \mathbf{U}_0$,得到

$$\tilde{\mathbf{m}}_p = \mathbf{V}_p\boldsymbol{\Lambda}_p^{-1}\mathbf{U}_p^T\mathbf{d} \tag{2.106}$$

所以算子 $\mathbf{A}$ 的广义逆算子为

$$\mathbf{A}^{-g}=\mathbf{V}_p\mathbf{\Lambda}_p^{-1}\mathbf{U}_p^T \tag{2.107}$$

姑且把它称为SVD广义逆。

求SVD广义逆归根到底是求本征值和本征矢量问题。

习题：已知线性系统 $\begin{cases} m_1+m_2=1 \\ m_3=2 \\ -m_3=1 \end{cases}$，求该系统的广义逆。写出数据在 $\mathbf{U}_p$ 和 $\mathbf{U}_0$ 空间中的分量。

2.9 SVD广义逆反演的意义

先考察第一种情况，$\mathbf{U}_0$和$\mathbf{V}_0$ 都不存在。这时 $p=m=n$，$\mathbf{U}_p$和$\mathbf{V}_p$ 是满秩的归一化正交矩阵，$\mathbf{U}_p^T\mathbf{U}_p=\mathbf{U}_p\mathbf{U}_p^T=\mathbf{V}_p^T\mathbf{V}_p=\mathbf{V}_p\mathbf{V}_p^T=\mathbf{I}$，广义逆为

$$\mathbf{A}^{-g}=\mathbf{V}_p\mathbf{\Lambda}_p^{-1}\mathbf{U}_p^T=(\mathbf{U}_p\mathbf{\Lambda}_p\mathbf{V}_p^T)^{-1}=\mathbf{A}^{-1} \tag{2.108}$$

线性方程组的解为

$$\tilde{\mathbf{m}}=\mathbf{A}^{-1}\mathbf{d} \tag{2.109}$$

解是完美的。

第二种情况，$\mathbf{V}_0$ 存在，但 $\mathbf{U}_0$ 不存在。方程系统是和谐的，$p=n<m$，$\mathbf{U}_p$ 满秩，$\mathbf{U}_p^T\mathbf{U}_p=\mathbf{U}_p\mathbf{U}_p^T=\mathbf{I}$，$\mathbf{V}_p^T\mathbf{V}_p=\mathbf{I}$，但 $\mathbf{V}_p\mathbf{V}_p^T\neq\mathbf{I}$，这时

$$\mathbf{A}\mathbf{A}^{-g}=\mathbf{U}_p\mathbf{\Lambda}_p\mathbf{V}_p^T\mathbf{V}_p\mathbf{\Lambda}_p^{-1}\mathbf{U}_p^T=\mathbf{I} \tag{2.110}$$

设 $(\mathbf{A}\mathbf{A}^T)^{-1}$ 存在，由（2.110）

$$\mathbf{A}\mathbf{A}^{-g}=\mathbf{A}\mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}$$

所以

$$\mathbf{A}^{-g}=\mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1} \tag{2.111}$$

方程系统的解为

$$\tilde{\mathbf{m}}=\mathbf{A}^{-g}\mathbf{d}=\mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}\mathbf{d} \tag{2.112}$$

与（2.19）比较可知，此时的解就是最小范数解。

*讨论：从图5讨论最小范数解

第三种情况，$\mathbf{U}_0$ 存在，$\mathbf{V}_0$ 不存在。此时方程是矛盾的超定系统，$p=m<n$，$\mathbf{V}_p$ 满秩，$\mathbf{V}_p^T\mathbf{V}_p=\mathbf{V}_p\mathbf{V}_p^T=\mathbf{I}$，$\mathbf{U}_p^T\mathbf{U}_p=\mathbf{I}$，但 $\mathbf{U}_p\mathbf{U}_p^T\neq\mathbf{I}$，这时

$$\mathbf{A}^{-g}\mathbf{A}=\mathbf{V}_p\mathbf{\Lambda}_p^{-1}\mathbf{U}_p^T\mathbf{U}_p\mathbf{\Lambda}_p\mathbf{V}_p^T=\mathbf{I} \tag{2.113}$$

设$(\mathbf{A}^T\mathbf{A})^{-1}$存在，由（2.113）

$$\mathbf{A}^{-g}\mathbf{A}=(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{A}$$

所以 $$\mathbf{A}^{-g}=(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T \tag{2.114}$$

方程系统的解为

$$\tilde{\mathbf{m}}=\mathbf{A}^{-g}\mathbf{d}=(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{d} \tag{2.115}$$

与（2.10）比较可知，此时的解就是最小二乘解。

第四种情况，$\mathbf{U}_0$存在，$\mathbf{V}_0$存在。此时$p<n$，m，$(\mathbf{A}^T\mathbf{A})^{-1}$与$(\mathbf{A}\mathbf{A}^T)^{-1}$都不存在。此种情况就是图5. 由图5知，不管哪种情况都是求方程$\mathbf{A}\mathbf{m}_p=\mathbf{d}_p$的解，所以

$$\mathbf{d}-\mathbf{A}\mathbf{m}_p=\mathbf{d}_0 \tag{2.116}$$

由于子空间$\mathbf{U}_0$与子空间$\mathbf{U}_p$相互垂直，$\mathbf{d}_0$是$\mathbf{U}_0$空间中的矢量，$\mathbf{A}\mathbf{m}_p$是$\mathbf{U}_p$空间中的分量，所以$\mathbf{d}_0$与$\mathbf{A}\mathbf{m}_p$垂直，范数$\|\mathbf{d}-\mathbf{A}\mathbf{m}_p\|$最小。图6使这种情况的图示。因此，这种情况下的SVD广义逆解仍然是最小二乘解，尽管$(\mathbf{A}^T\mathbf{A})^{-1}$与$(\mathbf{A}\mathbf{A}^T)^{-1}$都不存在。另外，$\mathbf{V}_0$的存在意味着解的不唯一，但$\mathbf{m}_p$不含$\mathbf{V}_0$分量，所以是最小范数解。

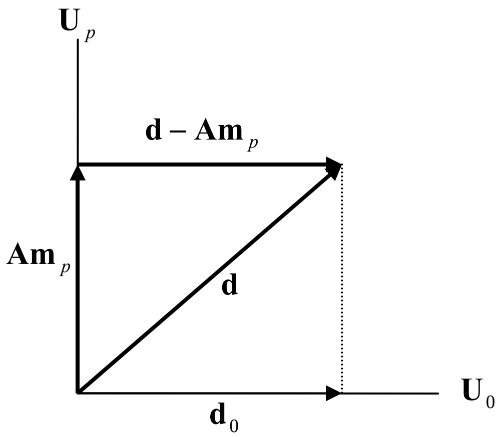

图 6　SVD 广义逆解的几何解释

2.10　SVD广义逆解的分辨与误差分析

由（1.4）知SVD广义逆的分辨矩阵为

$$\mathbf{R}=\mathbf{A}^{-g}\mathbf{A}=\mathbf{V}_p\mathbf{\Lambda}_p^{-1}\mathbf{U}_p^T\mathbf{U}_p\mathbf{\Lambda}_p\mathbf{V}_p^T=\mathbf{V}_p\mathbf{V}_p^T \qquad (2.117)$$

由（2.3）得到

$$\tilde{\mathbf{m}}=\mathbf{V}_p\mathbf{V}_p^T\mathbf{m}+\mathbf{A}^{-g}\boldsymbol{\varepsilon} \qquad (2.118)$$

当$\mathbf{V}_0$空间不存在时，$\mathbf{V}_p\mathbf{V}_p^T=\mathbf{I}$，不管$\mathbf{U}_0$是否存在解都是唯一的。当$\mathbf{V}_0$存在时，虽然解不唯一，但SVD广义逆只给出$\mathbf{V}_p$空间中的解，毫不理会$\mathbf{V}_0$空间的存在。这时的解是真模型的加权平均。由于$\mathbf{V}_0^T\tilde{\mathbf{m}}=\mathbf{0}$，$m-p$个约束方程，所以模型参数只有$p$个是独立的。如果模型参数是按真实物理量那样排列的，比如模型参数代表介质的速度，并按深度排列，那么分辨矩阵的形状就很重要。在不能完全分辨的情况下，我们总希望非零值集中在分辨矩阵的对角线附近，那么某一深度的速度就是真模型在同样深度附近的加权平均。另外，分辨矩阵的对角元素之和（矩阵的迹）等于p，我们可以把对角元素分成p组，每组的对角元素之和等于1，模型参数也相应的分成p组，那么每组的平均模型参数就是一个独立的估计值。

利用SVD广义逆还可以评价数据的质量。由于$\tilde{\mathbf{m}}=\mathbf{A}^{-g}\mathbf{d}$，则

$$\mathbf{d}_p=\mathbf{A}\tilde{\mathbf{m}}=\mathbf{A}\mathbf{A}^{-g}\mathbf{d}=\mathbf{U}_p\mathbf{U}_p^T\mathbf{d} \qquad (2.119)$$

矩阵$\mathbf{U}_p\mathbf{U}_p^T$为数据的分辨矩阵，当$\mathbf{U}_0$不存在时，$\mathbf{U}_p\mathbf{U}_p^T=\mathbf{I}$，理论数据与观测数据完全拟合。否则，理论数据就是观测数据的加权平均。由于$\mathbf{U}_0^T\mathbf{d}_p=\mathbf{0}$, $n-p$约束条件，所以理论数据只有p个分量是独立的。

以下我们来讨论模型的误差。由（2.5），模型的协方差矩阵为

$$\left\langle\Delta\tilde{\mathbf{m}}(\Delta\tilde{\mathbf{m}})^T\right\rangle=\mathbf{A}^{-g}\left\langle\Delta\mathbf{d}(\Delta\mathbf{d})^T\right\rangle\left(\mathbf{A}^{-g}\right)^T \qquad (2.120)$$

模型的方差

$$\sigma_{m_\alpha}^2=\sum_{\beta=1}\left(\mathbf{A}_{\alpha\beta}^{-g}\sigma_{d_\beta}\right)^2$$

假设数据是相互独立的，并且每个数据的方差都相同，为σ_d^2,那么（2.120）可简化为

$$\begin{aligned}\left\langle\Delta\tilde{\mathbf{m}}(\Delta\tilde{\mathbf{m}})^T\right\rangle&=\sigma_d^2\mathbf{A}^{-g}\left(\mathbf{A}^{-g}\right)^T\\&=\sigma_d^2\mathbf{V}_p\mathbf{\Lambda}_p^{-1}\mathbf{U}_p^T\mathbf{U}_p\mathbf{\Lambda}_p^{-1}\mathbf{V}_p^T\\&=\sigma_d^2\mathbf{V}_p\mathbf{\Lambda}_p^{-2}\mathbf{V}_p^T \qquad (2.121)\end{aligned}$$

由（2.121）可知，如果奇异值λ_α接近零，则模型的方差就会变得很大。为了减小模型的方差，可以将小的奇异值去除，并设定它们为零。但p值也相应的减少，这样就降低了模型和数据的分辨。实际工作中，要处理好分辨与误差的关系，鱼和熊掌不可兼得，在它们之间

要做出合理的权衡。

2.11　阻尼最小二乘法

在SVD广义逆解的分辨与误差分析中知道，如果出现接近零的奇异值，模型的方差就会变得很大。为了压制误差的放大，采用删除小的奇异值，令其为零的办法。这样虽然能够减小误差，但同时也降低了模型和数据的分辨。这种办法显得有点武断，理论上也缺乏严谨性。另外一种办法就是采用阻尼最小二乘法。其基本思想是压制小的奇异值，使误差控制在可接受的范围。实际上我们在2.3节混合问题中已经得到了这个结果，只不过那时我们没有把问题看得那么透彻。由（2.31），模型的解为

$$\tilde{\mathbf{m}} = \mathbf{A}^T(\mathbf{A}\mathbf{A}^T + \gamma\mathbf{I})^{-1}\mathbf{d}, \quad \gamma \ge 0 \qquad (2.122)$$

广义逆为

$$\mathbf{A}^{-g} = \mathbf{A}^T(\mathbf{A}\mathbf{A}^T + \gamma\mathbf{I})^{-1} \qquad (2.123)$$

把 $\mathbf{A}$ 的表达式(2.93)代入（2.123）得

$$\begin{aligned}\mathbf{A}^{-g} &= \mathbf{A}^T\left(\mathbf{U}_p(\mathbf{\Lambda}_p^2 + \gamma\mathbf{I})^{-1}\mathbf{U}_p^T + \mathbf{U}_0\gamma^{-2}\mathbf{U}_0^T\right) \\ &= \mathbf{V}_p\mathbf{\Lambda}_p\mathbf{U}_p^T\left(\mathbf{U}_p(\mathbf{\Lambda}_p^2 + \gamma\mathbf{I})^{-1}\mathbf{U}_p^T + \mathbf{U}_0\gamma^{-2}\mathbf{U}_0^T\right) \\ &= \mathbf{V}_p\mathbf{\Lambda}_p(\mathbf{\Lambda}_p^2 + \gamma\mathbf{I})^{-1}\mathbf{U}_p^T\end{aligned}$$

$$\mathbf{A}^{-g} = \mathbf{V}_p\frac{\mathbf{\Lambda}_p}{(\mathbf{\Lambda}_p^2 + \gamma\mathbf{I})}\mathbf{U}_p^T \qquad (2.124)$$

为阻尼最小二乘逆。

$$\tilde{\mathbf{m}} = \mathbf{V}_p\frac{\mathbf{\Lambda}_p}{(\mathbf{\Lambda}_p^2 + \gamma\mathbf{I})}\mathbf{U}_p^T\mathbf{d} \qquad (2.125)$$

为阻尼最小二乘的解。

设

$$\mathbf{d} = \mathbf{d}_p + \mathbf{d}_0 = \mathbf{U}_p\begin{pmatrix}\alpha_1\\ \vdots \\ \alpha_p\end{pmatrix} + \mathbf{U}_0\begin{pmatrix}\beta_1\\ \vdots \\ \alpha_{n-p}\end{pmatrix} = \mathbf{U}_p\boldsymbol{\alpha}_p + \mathbf{U}_0\boldsymbol{\beta}_p \qquad (2.126)$$

代入（1.125）得

$$\tilde{\mathbf{m}} = \mathbf{V}_p\frac{\mathbf{\Lambda}_p}{(\mathbf{\Lambda}_p^2 + \gamma\mathbf{I})}\boldsymbol{\alpha}_p \qquad (2,127)$$

说明，$\tilde{\mathbf{m}}$ 仍然是矢量。由于阻尼参数的任意性，一般情况下 $\tilde{\mathbf{m}}$ 不是最小范数解。容易证明，当 γ=0 时，$\tilde{\mathbf{m}}$ 是 $\mathbf{V}_p$ 空间中的最大范数解。

另外，

$$\mathbf{d} - \mathbf{A}\tilde{\mathbf{m}} = \mathbf{d} - \mathbf{A}\mathbf{V}_p\frac{\mathbf{\Lambda}_p}{(\mathbf{\Lambda}_p^2 + \gamma\mathbf{I})}\boldsymbol{\alpha}_p = \left(\mathbf{d}_p - \mathbf{U}_p\frac{\mathbf{\Lambda}_p^2}{(\mathbf{\Lambda}_p^2 + \gamma\mathbf{I})}\boldsymbol{\alpha}_p\right) + \mathbf{d}_0 \qquad (2.127)$$

由（2.127），当 $\gamma \ne 0$ 时，$\mathbf{d} - \mathbf{A}\tilde{\mathbf{m}} \ne \mathbf{d}_0$，说明 $\tilde{\mathbf{m}}$ 也不是最小二乘解。

根据推导（2.121）同样的假设，解的协方差矩阵为

$$\left\langle \Delta\tilde{\mathbf{m}}(\Delta\tilde{\mathbf{m}})^T \right\rangle = \sigma_d^2 \mathbf{V}_p \frac{\mathbf{\Lambda}_p^2}{(\mathbf{\Lambda}_p^2 + \gamma\mathbf{I})^2} \mathbf{V}_p^T \tag{2.128}$$

当奇异值接近零时，只要$\gamma \neq 0$，就能在一定的程度上压制误差的放大。

阻尼最小二乘准则的分辨矩阵为

$$\mathbf{R} = \mathbf{A}^{-g}\mathbf{A} = \mathbf{V}_p \frac{\mathbf{\Lambda}_p^2}{(\mathbf{\Lambda}_p^2 + \gamma\mathbf{I})} \mathbf{V}_p^T \tag{2.129}$$

分辨矩阵的迹为

$$\mathbf{R}\text{的迹}=\sum_{\alpha=1}^{p} \frac{\lambda_\alpha^2}{\lambda^2 + \gamma} \tag{2.130}$$

所以只要$\gamma > 0$，$\mathbf{R}$的迹 就小于 p，就会降低模型的分辨。

（习题：求分辨矩阵 $\mathbf{R}$ 的迹。）

实际问题中γ如何选取，带有一定的主观性。仍然是模型的分辨与误差之间权衡取舍。

2.12 SVD逆的协调性问题

令数据空间的变换为$\mathbf{d}' = \mathbf{Q}\mathbf{d}$，模型空间的变换为$\mathbf{m}' = \mathbf{S}\mathbf{m}$，并且设$\mathbf{Q}^{-1}, \mathbf{S}^{-1}$存在。原空间的本征值问题为

$$\begin{cases} \mathbf{A}\mathbf{v}_\alpha = \lambda_\alpha \mathbf{u}_\alpha \\ \mathbf{A}^T\mathbf{u}_\alpha = \lambda_\alpha \mathbf{v}_\alpha \end{cases} \tag{2.131}$$

由于$\mathbf{v}_\alpha$是模型空间中的矢量，$\mathbf{u}_\alpha$是数据空间中的矢量，所以

$$\mathbf{v}'_\alpha = \mathbf{S}\mathbf{v}_\alpha,\ \mathbf{u}' = \mathbf{Q}\mathbf{u} \tag{2.132}$$

代入（2.131）

$$\begin{cases} \mathbf{Q}\mathbf{A}\mathbf{S}^{-1}\mathbf{v}'_\alpha = \lambda_\alpha \mathbf{u}'_\alpha \\ \mathbf{S}\mathbf{A}^T\mathbf{Q}^{-1}\mathbf{u}'_\alpha = \lambda_\alpha \mathbf{v}'_\alpha \end{cases} \tag{2.133}$$

令$\mathbf{A}' = \mathbf{Q}\mathbf{A}\mathbf{S}^{-1}$，如果要求变换后的SVD逆协调，那么就必须令$(\mathbf{A}')^T = \mathbf{S}\mathbf{A}^T\mathbf{Q}^{-1}$，即

$$\mathbf{A} = (\mathbf{Q}^T\mathbf{Q})^{-1}\mathbf{A}\mathbf{S}^T\mathbf{S} \tag{2.134}$$

由（2.134）可知，除非 $\mathbf{Q}$ 和$\mathbf{S}$都是幺正变换，否则（2.134）不可能成立。幺正变换是空间的旋转，不会改变空间的标度。因此，一般情况下，SVD逆是不协调的。

习题：用方程组$\begin{cases} m_1 = 1 \\ m_2 = 2 \\ m_1 + m_2 = 2 \end{cases}$ 和 $\begin{cases} m_1 = 1 \\ m_2 = 2 \\ 2m_1 + 2m_2 = 4 \end{cases}$ 验证SVD逆不协调。

2.13 非线性问题中的迭代

以上探讨的都是线性方程组的反演方法。但是地球物理反演中的大部分问题都是非线性的。对于非线性问题，可采用线性迭代的办法，逐步逼近非线性问题的解。

设价值函数为

$$S = \|\mathbf{b} - \mathbf{f}(\mathbf{a})\|^2$$

$$= \left(b_i - f_i(\mathbf{a})\right)^2 \qquad (2.134)$$

其中 $b_i \ (i = 1, 2, \cdots n)$ 为观测数据，$a_i \ (i = 1, 2, \cdots m)$ 为模型参数。f_i 是关于模型 $\mathbf{a}$ 的非线性函数。更一般的情况，可以把 f_i 看成一个关于 $\mathbf{a}$ 的非线性算符。为讨论简单起见，我们把它看成一个关于 $\mathbf{a}$ 的函数。反演问题是求满足价值函数 S 极小值的模型 $\tilde{\mathbf{a}}$ 。(2.134)是 L_2 范数极小，实际问题中，可以根据不同的问题，提出不同的价值函数，然后求满足价值函数极小值的模型。设有一个初始模型 $\mathbf{a}^{(0)}$，首先对 f_i 进行线性化

$$f_i(\mathbf{a}) \approx f_i(\mathbf{a}^{(0)}) + \left.\frac{\partial f_i(\mathbf{a})}{\partial a_j}\right|_{\mathbf{a}=\mathbf{a}^{(0)}} (a_j - a_j^{(0)}) \qquad (2.135)$$

令 $d_i = b_i - f_i(\mathbf{a}^{(0)})$，$m_j = a_j - a_j^{(0)}$，$A_{ij} = \left.\dfrac{\partial f_i(\mathbf{a})}{\partial a_j}\right|_{\mathbf{a}=\mathbf{a}^{(0)}}$ (2.136)

将（2.135）代入（2.134）得

$$S \approx (d_i - A_{ij} m_j)^2 \qquad (2.137)$$

这是我们熟悉的最小二乘准则，根据（2.10）其解为

$$\tilde{\mathbf{m}} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{d} \qquad （2.138）$$

由于最初的问题是非线性的，要求出满足（2.134）的价值函数极小值的解，就必须迭代，逐步逼近。设由（2.138）第一次迭代的结果为 $\tilde{\mathbf{m}}^{(\mathbf{0})}$，根据（2.136）模型的第一次逼近则为

$$a_j^{(1)} = a_j^{(0)} + \tilde{m}_j^{(0)} \qquad (2.139)$$

第二次迭代，以 $\mathbf{a}^{(1)}$ 为初始模型，重复（2.135）－（2.138）的步骤，得到模型的二次逼近解。直到满足我们给定的收敛准则。

在实际问题中，以上的迭代过程可能不收敛，这取决于初始模型的选择、数据的质量等诸多因素。一个改进的办法是

$$a_j^{(1)} = a_j^{(0)} + \varepsilon \tilde{m}_j^{(0)}, \quad 0 < \varepsilon \le 1 \tag{2.140}$$

就是减小每一次迭代后修正的步长，使搜索变慢。实践证明，对有些问题即使ε取得很小，迭代仍然不收敛。

另一种完全不同的思路是采用梯度法或叫最速下降法。给定初始模型$\mathbf{a}^{(0)}$，求该点处评价函数的最速下降方向，并在该方向上寻找局部最小值，步步逼近最优解。$\mathbf{a}^{(0)}$点处的最速下降方向是评价函数的负梯度 $-\frac{\partial S}{\partial a_j}$，根据（2.134）得

$$\left.-\frac{\partial S}{\partial a_j}\right|_{\mathbf{a}=\mathbf{a}^{(0)}} = 2(b_i - f_i(\mathbf{a}^{(0)}))\left.\frac{\partial f_i(\mathbf{a})}{\partial a_j}\right|_{\mathbf{a}=\mathbf{a}^{(0)}} \tag{2.141}$$

$$= 2d_i A_{ij}$$

所以，评价函数的最速下降方向与$\mathbf{A}^T\mathbf{d}$平行。梯度法总是可以收敛的，但可能收敛的非常慢。另一方面，最小二乘迭代如果收敛，收敛的又过快，以致漏掉最佳解。在接近收敛时，最速下降方向$\mathbf{A}^T\mathbf{d}$与最小二乘迭代解$\tilde{\mathbf{m}}$几乎垂直。为了证明这点，作$\mathbf{A}^T\mathbf{d}$与$\tilde{\mathbf{m}}$的点乘

$$(\mathbf{A}^T\mathbf{d})\cdot\tilde{\mathbf{m}} = (\mathbf{d}^T\mathbf{U}_p\Lambda_p\mathbf{V}_p^T)\mathbf{V}_p\boldsymbol{\Lambda}_p^{-1}\mathbf{U}_p^T\mathbf{d}$$

$$= (\mathbf{U}_p^T\mathbf{d})^T\mathbf{U}_p^T\mathbf{d} \tag{2.142}$$

设 $\mathbf{d} = \mathbf{d}_p + \mathbf{d}_0 = \mathbf{U}_p\begin{pmatrix}\alpha_1\\ \vdots \\ \alpha_p\end{pmatrix} + \mathbf{U}_0\begin{pmatrix}\beta_1\\ \vdots \\ \alpha_{n-p}\end{pmatrix} = \mathbf{U}_p\boldsymbol{\alpha}_p + \mathbf{U}_0\boldsymbol{\beta}_p$，代入（2.142）得

$$(\mathbf{A}^T\mathbf{d})\cdot\tilde{\mathbf{m}} = \|\boldsymbol{\alpha}_p\|^2$$

根据（2.136）$\mathbf{d}$的定义，它是观测值与理论值的差，接近收敛时，$\mathbf{d}$的绝对值应该很小，$\boldsymbol{\alpha}_p$作为$\mathbf{d}_p$在$\mathbf{U}_p$空间中的坐标，其绝对值就更小。所以，最速下降方向$\mathbf{A}^T\mathbf{d}$与最小二乘迭代解$\tilde{\mathbf{m}}$几乎垂直。

如果能找到一个在它们二者之间的方向，那么迭代收敛的速度将适中。Marquardt(1963)指出了这一方向。那就是阻尼最小二乘法迭代的方向。由（2.27）

$$\tilde{\mathbf{m}} = (\mathbf{A}^T\mathbf{A} + \gamma\mathbf{I})^{-1}\mathbf{A}^T\mathbf{d} \tag{2.143}$$

当$\gamma = 0$时，（2.143）就是最小二乘解，当$\gamma \to \infty$时，$\tilde{\mathbf{m}} \to \gamma^{-1}\mathbf{A}^T\mathbf{d}$。所以，只要适当的选择阻尼系数，迭代就能迅速收敛。

习题：利用走时进行地震定位时常碰到求如下价值函数的极小值

$$\chi^2 = \sum_{\alpha=1}^{n} \left(\frac{t_\alpha - T_\alpha(\mathbf{a})}{\sigma_\alpha} \right)^2$$

其中t_α是第α个观测走时，T_α是第α个理论走时，$\mathbf{a} = (a_1, a_2, a_3, a_4)$的分量代表震源的空间坐标和发震时刻，在此问题中它们是待求的模型参数。把该问题线性化，并写出线性方程系统。

3. 非线性反演

对于线性离散系统，可以用以上介绍的SVD逆或阻尼最小二乘法求解。对于非线性问题，解决的方案是用线性迭代，逐步逼近非线性问题的解。一般来说，迭代是否收敛以及收敛到什么样的解与初始模型的选择密切相关。本章将介绍一种全新的方法，它不仅适用于线性系统的反演，同样适用于非线性系统。虽然对于非线性系统同样需要迭代，但与初始模型的选择的关系不像以前介绍的方法那么密切。对于连续性问题，处理方法与离散问题具有相同的思路。其解是完全和谐的。本方法由Tarantola和Vallete(1982)首先提出，并在地球物理反演中得到广泛的应用。我们把这种反演方法叫着Tarantola方法。

3.1 离散问题

基本假设：观测数据和模型参数是随机变量，它们满足高斯分布。

以前介绍的方法都是把模型参数和数据截然分开，这样处理隐含着一个假设，数据能够用模型参数显式的表示。对于更一般的情况，有时模型参数不能显式的表示数据，以前的方法就会遇到困难。为了解决这些困难，在此我们可以考虑数据和模型分量的联合空间，把数据与模型分量视为平等的地位。设联合空间的矢量为

$$\mathbf{x} = \begin{pmatrix} d_1 & d_2 & \cdots & d_n & m_1 & m_2 & \cdots & m_m \end{pmatrix}^T \qquad (3.1.1)$$

$\mathbf{x}$的各分量都被认为是随机变量，并设$\mathbf{x}$中的各分量满足高斯分布并存在先验均值$\mathbf{x}_0$。由于我们事先并不确切的知道$\mathbf{x}$的真正均值，所以$\mathbf{x}_0$也是随机变量序列。假设我们已经知道$\mathbf{x}_0$的协方差矩阵$\mathbf{C}_0$，称其为先验协方差矩阵。

实际问题中，观测数据往往是由一次实验获得的，我们并不真的知道这次观测数据的平均值，例如一次地震记录到的波形数据。但是，如果我们的观测仪器正常，再根据以往的观测经验，我们就有把握地说，观测数据在其平均值附近，或者说在观测误差之内。如果我们第一次看地震图，那是无论如何也无法判断其是正确的记录还是一条坏的记录（比如仪器出了问题，记录到的一张图），等到我们的经验积累到一定程度的时候，也许一眼就能判断出这条记录能不能用了。所谓经验，是我们过去读了很多图，也就相当于相同的实验重复做了很多次，才有了误差的标准，逐渐成为我们判断是非的准则。这些经验就成为我们判断将来事物的先验信息。既然我们并不确切的知道$\mathbf{x}$中数据分量的平均值，那么其均值就可以被看作随机变量，而观测数据就是$\mathbf{x}$中数据分量平均值的先验估计值。$\mathbf{x}$中的模型分量的先验均值取先验估计模型或叫初始模型的值。设观测数据序列为$\mathbf{d}_0 = (d_{01}, d_{02}, \cdots, d_{0n})$，先验模

型参数序列为 $\mathbf{m}_0=(m_{01}, m_{02}, \cdots, m_{0m})$，则 $\mathbf{x}$ 的先验均值为

$$\mathbf{x}_0=\begin{pmatrix}d_{01} & d_{02} & \cdots & d_{0n} & m_{01} & m_{02} & \cdots & m_{0m}\end{pmatrix}^T 。$$

先验信息中，数据的误差，数据间是否独立等是我们在反演前必须知道的，这些信息是在反演中得不到的。对于非线性问题，先验模型也非常重要。对于非线性问题，由于不是在全模型空间搜索，而是采取迭代的办法，这样就有可能迭代收敛到局部极小。不同的初始模型就有可能收敛到不同的结果。所以，要根据已有的信息，尽可能的给出物理上合理的初始模型。

离散问题广义最小二乘准则

正问题：　　$f_i(\mathbf{x})=0$，$(i=1,\cdots,n)$

$\mathbf{x}=(x_1, x_2, \cdots, x_{n+m})$ （随机变量序列）。

反问题：已知 $\mathbf{x}$ 的先验均值 $\mathbf{x}_0$（随机变量序列），以及关于 $\mathbf{x}_0$ 的先验协方差 $\mathbf{C}_0$，在约束条件 $f_i(\mathbf{x})=0$，$(i=1,\cdots,n)$ 下求满足高斯概率密度

$$\rho(\mathbf{x})=const\cdot\exp\left\{-\frac{1}{2}(\mathbf{x}-\mathbf{x}_0)^T\mathbf{C}_0^{-1}(\mathbf{x}-\mathbf{x}_0)\right\}$$

极大值的解 $\mathbf{x}=\tilde{\mathbf{x}}$ 。

正问题中 f_i 一般应理解为非线性算符，作用对象为 $\mathbf{x}$，函数只是它的特殊形式。以上准则就是在约束条件 $f_i(\mathbf{x})=0$，$(i=1,\cdots,n)$ 下求满足 $(\mathbf{x}-\mathbf{x}_0)^T\mathbf{C}_0^{-1}(\mathbf{x}-\mathbf{x}_0)$ 极小值的 $\mathbf{x}=\tilde{\mathbf{x}}$ 。这实际上是一定条件下的广义最小二乘准则。

定义价值函数

$$S=(\mathbf{x}-\mathbf{x}_0)^T\mathbf{C}_0^{-1}(\mathbf{x}-\mathbf{x}_0) \tag{3.1.2}$$

$$=(x_i-x_{0i})(\mathbf{C}_0^{-1})_{ij}(x_j-x_{0j})$$

根据拉格朗日条件极值的求法，定义

$$L=(x_i-x_{0i})(\mathbf{C}_0^{-1})_{ij}(x_j-x_{0j})+\lambda_i f_i(\mathbf{x})$$

则

$$\begin{cases}\dfrac{\partial L}{\partial x_k}=0\\ f_i(\mathbf{x})=0\end{cases} \tag{3.1.3}$$

即

$$
\begin{cases}
2(\mathbf{C}_0^{-1})_{ij}(x_j - x_{0j}) + \lambda_j \dfrac{\partial f_j(\mathbf{x})}{\partial x_i} = 0 \\
f_i(\mathbf{x}) = 0
\end{cases}
\tag{3.1.4}
$$

由（3.1.4）的第一式得

$$
(x_k - x_{0k}) = -\frac{1}{2}\lambda_j \frac{\partial f_j(\mathbf{x})}{\partial x_i} C_{0ik} \tag{3.1.5}
$$

令

$$
F_{ij}(\mathbf{x}) = \frac{\partial f_i(\mathbf{x})}{\partial x_j} \tag{3.1.6}
$$

由于 f_i 是算子，作用对象为 $\mathbf{x}$，所以 F_{ij} 也是算子，符号 $F_{ij}(\mathbf{x})$ 表示算子与 $\mathbf{x}$ 有关。

则

$$
(x_k - x_{0k}) = -\frac{1}{2}\lambda_j F_{ji}(\mathbf{x}) C_{0ik} \tag{3.1.7}
$$

$$
F_{lk}(\mathbf{x})(x_k - x_{0k}) = -\frac{1}{2} F_{lk}(\mathbf{x}) C_{0ki} F_{ij}^T(\mathbf{x}) \lambda_j \tag{3.1.8}
$$

用（3.1.8）减去（3.1.3）的第二式得

$$
F_{lk}(\mathbf{x})(x_k - x_{0k}) - f_l(\mathbf{x}) = -\frac{1}{2} F_{lk}(\mathbf{x}) C_{0ki} F_{ij}^T(\mathbf{x}) \lambda_j
$$

$$
\lambda_j = -2(\mathbf{F}(\mathbf{x})\mathbf{C}_0\mathbf{F}^T(\mathbf{x}))_{jl}^{-1}\left\{F_{lk}(\mathbf{x})(x_k - x_{0k}) - f_l(\mathbf{x})\right\} \tag{3.1.9}
$$

将（3.1.9）代入（3.1.5）得

$$
(x_k - x_{0k}) = C_{0ki} F_{ij}^T(\mathbf{x})\left(\mathbf{F}(\mathbf{x})\mathbf{C}_0\mathbf{F}^T(\mathbf{x})\right)_{jl}^{-1}\left\{F_{lm}(\mathbf{x})(x_m - x_{0m}) - f_l(\mathbf{x})\right\}
$$

即

$$
\mathbf{x} = \mathbf{x}_0 + \mathbf{C}_0\mathbf{F}^T(\mathbf{x})\left(\mathbf{F}(\mathbf{x})\mathbf{C}_0\mathbf{F}^T(\mathbf{x})\right)^{-1}\left\{\mathbf{F}(\mathbf{x})(\mathbf{x} - \mathbf{x}_0) - \mathbf{f}(\mathbf{x})\right\} \tag{3.1.20}
$$

值得提醒的是，（3.1.20）中的 $\mathbf{F}(\mathbf{x})$ 表示 $\mathbf{F}$ 是一个与 $\mathbf{x}$ 有关的算子，其定义为（3.1.6），如果 f_i 是非线性算子，一般来说 $\mathbf{F}$ 也是非线性算子。虽然（3.1.20）的等式两边都含有未知数 $\mathbf{x}$，但在非线性问题中它为我们提供了一个迭代公式。如果第 n 次迭代得到的解为 $\mathbf{x}^{(n)}$，（3.1.20）式的左边就给出第 $n+1$ 次的迭代解。于是有

$$
\mathbf{x}^{(n+1)} = \mathbf{x}_0 + \mathbf{C}_0\mathbf{F}^T(\mathbf{x}^{(n)})\left(\mathbf{F}(\mathbf{x}^{(n)})\mathbf{C}_0\mathbf{F}^T(\mathbf{x}^{(n)})\right)^{-1}\left\{\mathbf{F}(\mathbf{x}^{(n)})(\mathbf{x}^{(n)} - \mathbf{x}_0) - \mathbf{f}(\mathbf{x}^{(n)})\right\} \tag{3.1.21}
$$

如果是线性问题，即 $\mathbf{f}(\mathbf{x}) = \mathbf{F}\mathbf{x} = \mathbf{0}$ 其中算子 $\mathbf{F}$ 与 $\mathbf{x}$ 无关，（3.1.20）给出解

$$
\tilde{\mathbf{x}} = \mathbf{x}_0 - \mathbf{C}_0\mathbf{F}^T\left(\mathbf{F}\mathbf{C}_0\mathbf{F}^T\right)^{-1}\mathbf{F}\mathbf{x}_0 \tag{3.1.22}
$$

所以，对于线性问题无需迭代。

现在讨论协调性问题。设坐标变换为 $\mathbf{x}' = \mathbf{Q}\mathbf{x}$，并设 $\mathbf{Q}^{-1}$ 存在，则 $\mathbf{x} = \mathbf{Q}^{-1}\mathbf{x}'$。根据协方差的定义

$$\begin{aligned}\mathbf{C}_0 &= \langle(\mathbf{x}-\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0)^T\rangle \\ &= \mathbf{Q}^{-1}\langle(\mathbf{x}'-\mathbf{x}_0')(\mathbf{x}'-\mathbf{x}_0')^T\rangle\mathbf{Q}^{-T} \\ &= \mathbf{Q}^{-1}\mathbf{C}_0'\mathbf{Q}^{-T}\end{aligned}$$

代入价值函数（3.1.2）得

$$S = (\mathbf{x}-\mathbf{x}_0)^T\mathbf{C}_0^{-1}(\mathbf{x}-\mathbf{x}_0) = (\mathbf{x}'-\mathbf{x}_0')^T\mathbf{C}_0'^{-1}(\mathbf{x}'-\mathbf{x}_0')$$

价值函数的形式不变，所以（3.1.20）的解是谐调的。

（3.1.20）表达的是混合空间的结果，我们也可以把它表达为另一种形式。设

$$\begin{aligned}&\mathbf{x} = \begin{pmatrix}\mathbf{d}\\ \mathbf{m}\end{pmatrix}\\ &\mathbf{d} = \begin{pmatrix}x_1 & x_2 & \cdots & x_n\end{pmatrix}^T = \begin{pmatrix}d_1 & d_2 & \cdots & d_n\end{pmatrix}^T\\ &\mathbf{m} = \begin{pmatrix}x_{n+1} & x_{n+2} & \cdots & x_{n+m}\end{pmatrix}^T = \begin{pmatrix}m_1 & m_2 & \cdots & m_m\end{pmatrix}^T\\ &\mathbf{d}_0 = \begin{pmatrix}x_{01} & x_{02} & \cdots & x_{0n}\end{pmatrix}^T = \begin{pmatrix}d_{01} & d_{02} & \cdots & d_{0n}\end{pmatrix}^T\\ &\mathbf{m}_0 = \begin{pmatrix}x_{0_{n+1}} & x_{0n+2} & \cdots & x_{0n+m}\end{pmatrix}^T = \begin{pmatrix}m_{01} & m_{02} & \cdots & m_{0m}\end{pmatrix}^T\end{aligned} \qquad (3.1.23)$$

也就是把 $\mathbf{x}$ 分为理论数据和理论模型两个分量， $\mathbf{x}_0$ 分为观测数据和先验模型两个分量。先验协方差为

$$\mathbf{C}_0 = \begin{pmatrix}\mathbf{C}_0^{dd} & \mathbf{C}_0^{dm}\\ \mathbf{C}_0^{md} & \mathbf{C}_0^{mm}\end{pmatrix} \qquad (3.1.24)$$

其中 $\mathbf{C}_0^{dd}$ 为观测数据的协方差矩阵， $\mathbf{C}_0^{mm}$ 为先验模型的协方差矩阵， $\mathbf{C}_0^{md}$ 和 $\mathbf{C}_0^{dm}$ 为观测数据与先验模型交叉协方差矩阵。算子 $\mathbf{F}(\mathbf{x})$ 可表为

$$\mathbf{F}(\mathbf{x}) = \begin{pmatrix}\mathbf{F}^d(\mathbf{x}) & \mathbf{F}^m(\mathbf{x})\end{pmatrix} \qquad (3.1.25)$$

其中 $\mathbf{F}^d(\mathbf{x})$ 和 $\mathbf{F}^m(\mathbf{x})$ 的分量分别定义为

$$F_{ij}^d = \frac{\partial f_i(\mathbf{x})}{\partial d_j}, \qquad F_{ij}^m = \frac{\partial f_i(\mathbf{x})}{\partial m_j} \qquad (3.1.26)$$

令 $\quad \mathbf{S}(\mathbf{x}) = \mathbf{F}(\mathbf{x})\mathbf{C}_0\mathbf{F}^T(\mathbf{x})$

$$=\left(\mathbf{F}^{d}(\mathbf{x}) \quad \mathbf{F}^{m}(\mathbf{x})\right)\begin{pmatrix}\mathbf{C}_0^{dd} & \mathbf{C}_0^{dm} \\ \mathbf{C}_0^{md} & \mathbf{C}_0^{mm}\end{pmatrix}\begin{pmatrix}\left(\mathbf{F}^{d}(\mathbf{x})\right)^T \\ \left(\mathbf{F}^{m}(\mathbf{x})\right)^T\end{pmatrix} \qquad 3.1.27）$$

把（3.123）－（3.125）代入（3.1.20）

$$\begin{pmatrix}\mathbf{d} \\ \mathbf{m}\end{pmatrix}=\begin{pmatrix}\mathbf{d}_0 \\ \mathbf{m}_0\end{pmatrix}+\begin{pmatrix}\mathbf{C}_0^{dd} & \mathbf{C}_0^{dm} \\ \mathbf{C}_0^{md} & \mathbf{C}_0^{mm}\end{pmatrix}\begin{pmatrix}\left(\mathbf{F}^{d}(\mathbf{x})\right)^T \\ \left(\mathbf{F}^{m}(\mathbf{x})\right)^T\end{pmatrix}\mathbf{S}^{-1}(\mathbf{x})\left\{\left(\mathbf{F}^{d}(\mathbf{x}) \quad \mathbf{F}^{m}(\mathbf{x})\right)\begin{pmatrix}\mathbf{d}-\mathbf{d}_0 \\ \mathbf{m}-\mathbf{m}_0\end{pmatrix}-\mathbf{f}(\mathbf{x})\right\}$$

所以

$$\begin{cases}\mathbf{d}=\mathbf{d}_0+\left(\mathbf{C}_0^{dd}\left(\mathbf{F}^{d}(\mathbf{x})\right)^T+\mathbf{C}_0^{dm}\left(\mathbf{F}^{m}(\mathbf{x})\right)^T\right)\mathbf{S}^{-1}(\mathbf{x})\left\{\left(\mathbf{F}^{d}(\mathbf{x}) \quad \mathbf{F}^{m}(\mathbf{x})\right)\begin{pmatrix}\mathbf{d}-\mathbf{d}_0 \\ \mathbf{m}-\mathbf{m}_0\end{pmatrix}-\mathbf{f}(\mathbf{x})\right\} \\ \mathbf{m}=\mathbf{m}_0+\left(\mathbf{C}_0^{md}\left(\mathbf{F}^{d}(\mathbf{x})\right)^T+\mathbf{C}_0^{mm}\left(\mathbf{F}^{m}(\mathbf{x})\right)^T\right)\mathbf{S}^{-1}(\mathbf{x})\left\{\left(\mathbf{F}^{d}(\mathbf{x}) \quad \mathbf{F}^{m}(\mathbf{x})\right)\begin{pmatrix}\mathbf{d}-\mathbf{d}_0 \\ \mathbf{m}-\mathbf{m}_0\end{pmatrix}-\mathbf{f}(\mathbf{x})\right\}\end{cases}$$

（3.1.28)

对于非线性问题，需要迭代。解释与（3.1.21）相同。

对于线性问题

$$\tilde{\mathbf{d}}=\mathbf{d}_0-\left(\mathbf{C}_0^{dd}\left(\mathbf{F}^{d}\right)^T+\mathbf{C}_0^{dm}\left(\mathbf{F}^{m}\right)^T\right)\mathbf{S}^{-1}\left(\mathbf{F}^{d}\mathbf{d}_0+\mathbf{F}^{m}\mathbf{m}_0\right) \qquad (3.1.29)$$

$$\tilde{\mathbf{m}}=\mathbf{m}_0-\left(\mathbf{C}_0^{md}\left(\mathbf{F}^{d}\right)^T+\mathbf{C}_0^{mm}\left(\mathbf{F}^{m}\right)^T\right)\mathbf{S}^{-1}\left(\mathbf{F}^{d}\mathbf{d}_0+\mathbf{F}^{m}\mathbf{m}_0\right) \qquad (3.1.30)$$

$$\mathbf{S}=\left(\mathbf{F}^{d} \quad \mathbf{F}^{m}\right)\begin{pmatrix}\mathbf{C}_0^{dd} & \mathbf{C}_0^{dm} \\ \mathbf{C}_0^{md} & \mathbf{C}_0^{mm}\end{pmatrix}\begin{pmatrix}\left(\mathbf{F}^{d}\right)^T \\ \left(\mathbf{F}^{m}\right)^T\end{pmatrix}$$

从（3.1.29)（3.1.30）可知，对于线性问题，理论数据和理论模型之间完全解耦。

以上讨论的是最一般的情况，给出的解是（3.1.27）－（3.1.29)。在实际的地球物理反演问题中，观测数据和先验模型是相互独立的，并且多数情况下，理论数据可用模型参数显式表示。下面就几种常见的情况进行讨论。

1) 观测数据与先验模型参数相互独立

在此情况下，$\mathbf{C}_0^{dm}=\mathbf{C}_0^{md}=\mathbf{0}$，

$$\mathbf{S}(\mathbf{x})=\mathbf{F}^{d}(\mathbf{x})\mathbf{C}_0^{dd}\left(\mathbf{F}^{d}(\mathbf{x})\right)^T+\mathbf{F}^{m}(\mathbf{x})\mathbf{C}_0^{mm}\left(\mathbf{F}^{m}(\mathbf{x})\right)^T \qquad （3.1.31）$$

$$\begin{cases}\mathbf{d}^{(n+1)}=\mathbf{d}_0+\mathbf{C}_0^{dd}\left(\mathbf{F}^{d}(\mathbf{x}^{(n)})\right)^T\mathbf{S}^{-1}(\mathbf{x}^{(n)})\left\{\left(\mathbf{F}^{d}(\mathbf{x}^{(n)}) \quad \mathbf{F}^{m}(\mathbf{x}^{(n)})\right)\begin{pmatrix}\mathbf{d}-\mathbf{d}_0 \\ \mathbf{m}-\mathbf{m}_0\end{pmatrix}-\mathbf{f}(\mathbf{x}^{(n)})\right\} \\ \mathbf{m}^{(n+1)}=\mathbf{m}_0+\mathbf{C}_0^{mm}\left(\mathbf{F}^{m}(\mathbf{x}^{(n)})\right)^T\mathbf{S}^{-1}(\mathbf{x}^{(n)})\left\{\left(\mathbf{F}^{d}(\mathbf{x}^{(n)}) \quad \mathbf{F}^{m}(\mathbf{x}^{(n)})\right)\begin{pmatrix}\mathbf{d}-\mathbf{d}_0 \\ \mathbf{m}-\mathbf{m}_0\end{pmatrix}-\mathbf{f}(\mathbf{x}^{(n)})\right\}\end{cases}$$

（3.1.32）

对线性问题

正问题：$\mathbf{f}(\mathbf{d},\mathbf{m})=\mathbf{F}\begin{pmatrix}\mathbf{d}\\ \mathbf{m}\end{pmatrix}=\mathbf{0}$，算子 $\mathbf{F}$ 不含 $\mathbf{d},\mathbf{m}$ 的分量。

即 $\mathbf{F}^d\mathbf{d}+\mathbf{F}^m\mathbf{m}=\mathbf{0}$

反问题解：

$$\mathbf{S}(\mathbf{x})=\mathbf{F}^d\mathbf{C}_0^{dd}\left(\mathbf{F}^d\right)^T+\mathbf{F}^m\mathbf{C}_0^{mm}\left(\mathbf{F}^m\right)^T$$

$$\tilde{\mathbf{d}}=\mathbf{d}_0-\mathbf{C}_0^{dd}\left(\mathbf{F}^d\right)^T\mathbf{S}^{-1}\left(\mathbf{F}^d\mathbf{d}_0+\mathbf{F}^m\mathbf{m}_0\right) \qquad (3.1.33)$$

$$\tilde{\mathbf{m}}=\mathbf{m}_0-\mathbf{C}_0^{mm}\left(\mathbf{F}^m\right)^T\mathbf{S}^{-1}\left(\mathbf{F}^d\mathbf{d}_0+\mathbf{F}^m\mathbf{m}_0\right) \qquad (3.1.34)$$

2) 观测数据与先验模型参数相互独立，数据可用模型参数显式表示

在此情况下，$\mathbf{C}_0^{dm}=\mathbf{C}_0^{md}=\mathbf{0}$，正演形式 $\mathbf{f}(\mathbf{d},\mathbf{m})=\mathbf{d}-\mathbf{g}(\mathbf{m})=\mathbf{0}$，即

$$\mathbf{d}=\mathbf{g}(\mathbf{m}) \text{ 或 } d_i=g_i(\mathbf{m}),\ (i=1,2,\cdots,n) \qquad (3.1.35)$$

根据定义（3.1.26）推得

$$\mathbf{F}^d(\mathbf{x})=\mathbf{I},\ \mathbf{F}^m(\mathbf{x})=-\mathbf{G}(\mathbf{m}),\ G_{ij}=\frac{\partial g_i(\mathbf{m})}{\partial m_j} \qquad (3.1.36)$$

将（3.1.35）（3.1.36）代入（3.1.31）—（3.1.34）得

$$\mathbf{S}(\mathbf{x})=\mathbf{C}_0^{dd}+\mathbf{G}(\mathbf{m})\mathbf{C}_0^{mm}\mathbf{G}^T(\mathbf{m}) \qquad (3.1.37)$$

$$\begin{cases}\mathbf{d}^{(n+1)}=\mathbf{d}_0-\mathbf{C}_0^{dd}\mathbf{S}^{-1}(\mathbf{m}^{(n)})\left\{(\mathbf{d}_0-\mathbf{g}(\mathbf{m}^{(n)}))+\mathbf{G}(\mathbf{m}^{(n)})(\mathbf{m}^{(n)}-\mathbf{m}_0)\right\}\\ \mathbf{m}^{(n+1)}=\mathbf{m}_0+\mathbf{C}_0^{mm}\mathbf{G}^T(\mathbf{m}^{(n)})\mathbf{S}^{-1}(\mathbf{m}^{(n)})\left\{(\mathbf{d}_0-\mathbf{g}(\mathbf{m}^{(n)}))+\mathbf{G}(\mathbf{m}^{(n)})(\mathbf{m}^{(n)}-\mathbf{m}_0)\right\}\end{cases} \qquad (3.1.38)$$

对于线性问题

由 $\mathbf{d}-\mathbf{G}\mathbf{m}=\mathbf{F}\begin{pmatrix}\mathbf{d}\\ \mathbf{m}\end{pmatrix}=\mathbf{0}$，推得 $\mathbf{F}=(\mathbf{I}\quad -\mathbf{G})$，其中算子 $\mathbf{G}$ 不含 $\mathbf{d},\mathbf{m}$ 的分量。所以，

$$\mathbf{F}^d(\mathbf{x})=\mathbf{I},\ \mathbf{F}^m(\mathbf{x})=-\mathbf{G}$$

正问题：$\mathbf{d}=\mathbf{G}\mathbf{m}$

反问题解：

$$\mathbf{S}=\mathbf{G}\mathbf{C}_0^{mm}\mathbf{G}^T+\mathbf{C}_0^{dd} \qquad (3.1.39)$$

$$\tilde{\mathbf{m}}=\mathbf{m}_0+\mathbf{C}_0^{mm}\mathbf{G}^T\mathbf{S}^{-1}\left(\mathbf{d}_0-\mathbf{G}\mathbf{m}_0\right) \qquad (3.1.40)$$

$$\tilde{\mathbf{d}}=\mathbf{d}_0-\mathbf{C}_0^{dd}\mathbf{S}^{-1}\left(\mathbf{d}_0-\mathbf{G}\mathbf{m}_0\right)=\mathbf{G}\tilde{\mathbf{m}} \qquad (3.1.41)$$

利用等式

$$\mathbf{C}_0^{mm}\mathbf{G}^T(\mathbf{G}\mathbf{C}_0^{mm}\mathbf{G}^T+\mathbf{C}_0^{dd})^{-1}=\left(\mathbf{G}^T(\mathbf{C}_0^{dd})^{-1}\mathbf{G}+(\mathbf{C}_0^{mm})^{-1}\right)^{-1}\mathbf{G}^T(\mathbf{C}_0^{dd})^{-1} \quad (3.1.42$$

（3.1.40）可写为另一形式

$$\tilde{\mathbf{m}}=\mathbf{m}_0+\left(\mathbf{G}^T(\mathbf{C}_0^{dd})^{-1}\mathbf{G}+(\mathbf{C}_0^{mm})^{-1}\right)^{-1}\mathbf{G}^T(\mathbf{C}_0^{dd})^{-1}\left(\mathbf{d}_0-\mathbf{G}\mathbf{m}_0\right) \quad （3.1.43）$$

如果令 $\mathbf{W}_d=(\mathbf{C}_0^{dd})^{-1}$，$\gamma\mathbf{W}_m=(\mathbf{C}_0^{mm})^{-1}$，$\mathbf{G}=\mathbf{A},\mathbf{m}_0=\mathbf{0}$，则（3.1.43）与（2.61）相同。相当于谐调的阻尼最小二乘解。
（课堂讨论：初始模型对解的影响。）

由于线性问题无需迭代，初始模型不是必须给出的，这对所有的线性问题都成立。所以，对所有的线性问题都可以设 $\mathbf{m}_0=\mathbf{0}$，但模型参数的先验协方差一般是要给出的。否则，对模型范数的要求将改变，反演结果可能是条件病态的（很小的观测数据误差引起很大的模型参数误差，参照阻尼最小二乘法的讨论）。对于非线性问题，由于 $\mathbf{m}_0$ 是迭代的起点，所以必须给出。

3） 无观测数据

无观测数据是反演的特例，在地球物理反演中一般不会遇到。但为了比较系统的讨论广义最小二乘法反演方法的特点，我们不妨简单讨论这一有趣的问题。实际上，这是方程的求根问题。

在此情况下，$\mathbf{C}_0^{dm}=\mathbf{C}_0^{md}=\mathbf{C}_0^{dd}=\mathbf{0}$，根据（3.1.31）（3.1.32）可知

正问题：$\mathbf{f}(\mathbf{m})=0$

反问题解：

$$\mathbf{S}(\mathbf{x})=\mathbf{F}^m(\mathbf{m})\mathbf{C}_0^{mm}\left(\mathbf{F}^m(\mathbf{m})\right)^T，\text{其中}\quad F_{ij}^m=\frac{\partial f_i(\mathbf{m})}{\partial m_j} \quad （3.1.44）$$

$$\mathbf{m}^{(n+1)}=\mathbf{m}_0+\mathbf{C}_0^{mm}\left(\mathbf{F}^m(\mathbf{m}^{(n)})\right)^T\mathbf{S}^{-1}(\mathbf{m}^{(n)})\left\{\mathbf{F}^m(\mathbf{m}^{(n)})(\mathbf{m}^{(n)}-\mathbf{m}_0)-\mathbf{f}(\mathbf{m}^{(n)})\right\} \quad （3.1.45）$$

4） 分辨和协方差

直接讨论非线性系统迭代结果的分辨和协方差有一定的困难。理论上不是不能做，但实际上我们没有必要那样做。如果一个非线性问题迭代收敛，就意味着收敛前的几次迭代结果相互之间相差很小。如果以这些点中的任意一点对非线性方程进行泰勒展开，展开到一级项就应该有足够的精度，非线性问题就被线性化了。所以，收敛前的几次非线性迭代就相当于线性迭代，其误差是二级以上的小量。因此，如果非线性迭代收敛，其结果就可以用线性系统中的评估方法对其进行评估。我们还是从一般的问题入手。线性系统的正问题为

$$\mathbf{f}(\mathbf{x})=\mathbf{F}\mathbf{x}=\mathbf{0} \quad （3.1.46）$$

其中算子 $\mathbf{F}$ 与 $\mathbf{x}$ 无关。根据（3.1.20），其反问题的解为

$$\mathbf{x} = \mathbf{x}_0 - \mathbf{C}_0\mathbf{F}^T\left(\mathbf{F}\mathbf{C}_0\mathbf{F}^T\right)^{-1}\mathbf{F}\mathbf{x}_0 \tag{3.1.47}$$

下角标形式为

$$x_i = x_{0i} - C_{0ij}F_{kj}(\mathbf{F}\mathbf{C}_0\mathbf{F}^T)^{-1}_{kl}F_{lm}x_{0m}$$

$$= \delta_{im}x_{0m} - C_{0ij}F_{kj}(\mathbf{F}\mathbf{C}_0\mathbf{F}^T)^{-1}_{kl}F_{lm}x_{0m}$$

$$= \left(\delta_{im} - C_{0ij}F_{kj}(\mathbf{F}\mathbf{C}_0\mathbf{F}^T)^{-1}_{kl}F_{lm}\right)x_{0m}$$

$$= B_{im}x_{0m} \qquad （3.1.48）$$

其中 $$B_{im} = \delta_{im} - C_{0ij}F_{kj}(\mathbf{F}\mathbf{C}_0\mathbf{F}^T)^{-1}_{kl}F_{lm}$$

即 $$\mathbf{B} = \mathbf{I} - \mathbf{C}_0\mathbf{F}^T(\mathbf{F}\mathbf{C}_0\mathbf{F}^T)^{-1}\mathbf{F} \qquad （3.1.49）$$

由（3.1.48）可知，矩阵 $\mathbf{B}$ 是数据的分辨矩阵。我们已经有了关于 $\mathbf{x}_0$ 的先验协方差，现在来求关于 $\mathbf{x}$ 的协方差，被称为后验协方差。根据协方差的定义

$$C_{ij} = \langle(x_i - \langle x_i\rangle)(x_j - \langle x_j\rangle)\rangle$$

由（3.1.48）得

$$C_{ij} = \langle(B_{im}x_{0m} - \langle B_{im}x_{0m}\rangle)(B_{jn}x_{0n} - \langle B_{jn}x_{0n}\rangle)\rangle$$

$$= B_{im}\langle(x_{0m} - \langle x_{0m}\rangle)(x_{0n} - \langle x_{0n}\rangle)\rangle B_{jn}$$

$$= B_{im}C_{0mn}B_{jn}$$

即 $$\mathbf{C} = \mathbf{B}\mathbf{C}_0\mathbf{B}^T$$

利用（3.1.49）化简为

$$\mathbf{C} = \mathbf{B}\mathbf{C}_0 \qquad （3.1.50）$$

我们讨论的是混合空间的后验协方差矩阵，它包含了后验的数据协方差和后验的模型协方差，根据这些协方差我们可以评估模型的误差和分辨。

因为

$$\mathbf{C} = \begin{pmatrix} \mathbf{C}^{dd} & \mathbf{C}^{dm} \\ \mathbf{C}^{md} & \mathbf{C}^{mm} \end{pmatrix} = \begin{pmatrix} \mathbf{C}_0^{dd} & \mathbf{C}_0^{dm} \\ \mathbf{C}_0^{md} & \mathbf{C}_0^{mm} \end{pmatrix} - \begin{pmatrix} \mathbf{C}_0^{dd} & \mathbf{C}_0^{dm} \\ \mathbf{C}_0^{md} & \mathbf{C}_0^{mm} \end{pmatrix}\begin{pmatrix} (\mathbf{F}^d)^T \\ (\mathbf{F}^m)^T \end{pmatrix}\mathbf{S}^{-1}\begin{pmatrix} \mathbf{F}^d & \mathbf{F}^m \end{pmatrix}\begin{pmatrix} \mathbf{C}_0^{dd} & \mathbf{C}_0^{dm} \\ \mathbf{C}_0^{md} & \mathbf{C}_0^{mm} \end{pmatrix}$$

所以

$$\mathbf{C}^{dd} = \mathbf{C}_0^{dd} - (\mathbf{C}_0^{dd}(\mathbf{F}^d)^T + \mathbf{C}_0^{dm}(\mathbf{F}^m)^T)\mathbf{S}^{-1}(\mathbf{F}^d\mathbf{C}_0^{dd} + \mathbf{F}^m\mathbf{C}_0^{md}) \qquad （3.1.51）$$

$$\mathbf{C}^{mm} = \mathbf{C}_0^{mm} - (\mathbf{C}_0^{md}(\mathbf{F}^d)^T + \mathbf{C}_0^{mm}(\mathbf{F}^m)^T)\mathbf{S}^{-1}(\mathbf{F}^d\mathbf{C}_0^{dm} + \mathbf{F}^m\mathbf{C}_0^{mm}) \qquad （3.1.52）$$

$$\mathbf{C}^{dm} = \mathbf{C}_0^{dm} - (\mathbf{C}_0^{dd}(\mathbf{F}^d)^T + \mathbf{C}_0^{dm}(\mathbf{F}^m)^T)\mathbf{S}^{-1}(\mathbf{F}^d\mathbf{C}_0^{dm} + \mathbf{F}^m\mathbf{C}_0^{mm}) \qquad （3.1.53）$$

$$\mathbf{C}^{md}=\mathbf{C}_0^{md}-(\mathbf{C}_0^{md}(\mathbf{F}^d)^T+\mathbf{C}_0^{mm}(\mathbf{F}^m)^T)\mathbf{S}^{-1}(\mathbf{F}^d\mathbf{C}_0^{dd}+\mathbf{F}^m\mathbf{C}_0^{md}) \quad (3.1.54)$$

$$\mathbf{S}=\begin{pmatrix}\mathbf{F}^d & \mathbf{F}^m\end{pmatrix}\begin{pmatrix}\mathbf{C}_0^{dd} & \mathbf{C}_0^{dm}\\ \mathbf{C}_0^{md} & \mathbf{C}_0^{mm}\end{pmatrix}\begin{pmatrix}(\mathbf{F}^d)^T\\ (\mathbf{F}^m)^T\end{pmatrix}$$

我们还可以定义一个与 $\mathbf{B}$ 等价的分辨矩阵

$$\mathbf{A}=\mathbf{C}_0\mathbf{F}^T(\mathbf{F}\mathbf{C}_0\mathbf{F}^T)^{-1}\mathbf{F} \quad (3.1.55)$$

其越接近单位矩阵，分辨就越高。以下讨论中我们会看到它的实际意义。对于非线性系统，如果迭代收敛，把最后一次计算得到的 $\mathbf{F}$ 代入以上各式中，就可以对计算模型进行评估。

设观测数据与先验模型参数相互独立，数据可用模型参数显式表示。则 $\mathbf{C}_0^{dm}=\mathbf{0}$，$\mathbf{C}_0^{md}=\mathbf{0}$，$\mathbf{d}=\mathbf{G}\mathbf{m}$，$\mathbf{F}=(\mathbf{F}^d \quad \mathbf{F}^m)$，$\mathbf{F}^d=\mathbf{I}$，$\mathbf{F}^m=-\mathbf{G}$，代入（3.1.51）－（3.1.55）得

$$\mathbf{S}=\mathbf{G}\mathbf{C}_0^{mm}\mathbf{G}^T+\mathbf{C}_0^{dd} \quad (3.1.56)$$

$$\mathbf{C}^{dd}=\mathbf{C}_0^{dd}-\mathbf{C}_0^{dd}\mathbf{S}^{-1}\mathbf{C}_0^{dd}=(\mathbf{I}-\mathbf{A}^{dd})\mathbf{C}_0^{dd} \quad (3.1.57)$$

$$\mathbf{C}^{mm}=\mathbf{C}_0^{mm}-\mathbf{C}_0^{mm}\mathbf{G}^T\mathbf{S}^{-1}\mathbf{G}\mathbf{C}_0^{mm}=(\mathbf{I}-\mathbf{A}^{mm})\mathbf{C}_0^{mm} \quad (3.1.58)$$

$$\mathbf{C}^{dm}=\mathbf{C}_0^{dd}\mathbf{S}^{-1}\mathbf{G}\mathbf{C}_0^{mm}=\mathbf{A}^{dm}\mathbf{C}_0^{mm} \quad (3.1.59)$$

$$\mathbf{C}^{md}=\mathbf{C}_0^{mm}\mathbf{G}^T\mathbf{S}^{-1}\mathbf{C}_0^{dd}=\mathbf{A}^{md}\mathbf{C}_0^{dd} \quad (3.1.60)$$

$$\mathbf{A}=\begin{pmatrix}\mathbf{A}^{dd} & \mathbf{A}^{dm}\\ \mathbf{A}^{md} & \mathbf{A}^{mm}\end{pmatrix} \quad (3.1.61)$$

我们来考察矩阵 $\mathbf{A}$ 各分量的意义。由（3.1.40）知，线性系统的解为

$$\tilde{\mathbf{m}}=\mathbf{m}_0+\mathbf{C}_0^{mm}\mathbf{G}^T\ \mathbf{S}^{-1}(\mathbf{d}_0-\mathbf{G}\mathbf{m}_0) \quad (3.1.62)$$

设真模型 $\mathbf{m}^r$ 是线性方程 $\mathbf{d}_0=\mathbf{G}\mathbf{m}+\boldsymbol{\varepsilon}$ 的解，$\boldsymbol{\varepsilon}$ 为误差项。则

$$\begin{aligned}\tilde{\mathbf{m}}-\mathbf{m}_0&=\mathbf{C}_0^{mm}\mathbf{G}^T\ \mathbf{S}^{-1}(\mathbf{G}\mathbf{m}^r-\mathbf{G}\mathbf{m}_0+\boldsymbol{\varepsilon})\\ &=\mathbf{C}_0^{mm}\mathbf{G}^T\ \mathbf{S}^{-1}\mathbf{G}(\mathbf{m}^r-\mathbf{m}_0)+\mathbf{C}_0^{mm}\mathbf{G}^T\ \mathbf{S}^{-1}\boldsymbol{\varepsilon}\\ &=\mathbf{A}^{mm}(\mathbf{m}^r-\mathbf{m}_0)+\mathbf{A}^{md}\boldsymbol{\varepsilon} \quad (3.1.63)\end{aligned}$$

(3.1.63)右边的第二项表示数据误差对解的影响，算子 $\mathbf{A}^{md}$ 把数据误差传递给模型参数。不妨把算子 $\mathbf{A}^{md}$ 称为误差传播算子。（3.1.63）右边的第一项表明反演模型是真模型的线性组

合。如果 $\mathbf{A}^{mm}=\mathbf{I}$，则反演模型 $\tilde{\mathbf{m}}$ 完全分辨。所以，算子 $\mathbf{A}^{mm}$ 是模型的分辨矩阵。

由（3.1.41）推知

$$
\begin{aligned}
\tilde{\mathbf{d}}-\mathbf{d}_0 &= \mathbf{C}_0^{dd}\mathbf{S}^{-1}\left(\mathbf{G}\mathbf{m}_0-\mathbf{d}_0\right) \\
&= \mathbf{C}_0^{dd}\mathbf{S}^{-1}\left(\mathbf{G}(\mathbf{m}^r-\mathbf{m}^r+\mathbf{m}_0)-\mathbf{d}_0\right) \\
&= \mathbf{C}_0^{dd}\mathbf{S}^{-1}\left(\mathbf{G}\mathbf{m}^r-\mathbf{d}_0+\mathbf{G}(\mathbf{m}_0-\mathbf{m}^r)\right) \\
&= \mathbf{C}_0^{dd}\mathbf{S}^{-1}\left(\mathbf{d}^r-\mathbf{d}_0)+\mathbf{C}_0^{dd}\mathbf{S}^{-1}\mathbf{G}(\mathbf{m}_0-\mathbf{m}^r)\right) \\
&= \mathbf{C}_0^{dd}\mathbf{S}^{-1}\left(\mathbf{d}^r-\mathbf{d}_0)+\mathbf{C}_0^{dd}\mathbf{S}^{-1}\mathbf{G}(\mathbf{m}_0-\mathbf{m}^r)\right) \\
&= \mathbf{A}^{dd}(\mathbf{d}^r-\mathbf{d}_0)+\mathbf{A}^{dm}(\mathbf{m}_0-\mathbf{m}^r) \qquad (3.1.64)
\end{aligned}
$$

其中 $\mathbf{d}^r=\mathbf{G}\mathbf{m}^r$ 为无误差的真数据。(3.1.64)表明，如果 $\mathbf{A}^{dd}=\mathbf{I}$，则计算数据完全分辨。所以 $\mathbf{A}^{dd}$ 是数据分辨矩阵。算子 $\mathbf{A}^{dm}$ 反映先验模型误差对计算数据的影响。利用（3.1.56）可以把（3.1.57）－（3.1.60）表为另一种形式

$$\mathbf{C}^{mm}=\mathbf{C}_0^{mm}-\mathbf{C}_0^{mm}\mathbf{G}^T\mathbf{S}^{-1}\mathbf{G}\mathbf{C}_0^{mm}=(\mathbf{I}-\mathbf{A}^{mm})\mathbf{C}_0^{mm}$$

$$\mathbf{C}^{dd}=\mathbf{G}\mathbf{C}^{mm}\mathbf{G}^T \qquad (3.1.65)$$

$$\mathbf{C}^{dm}=\mathbf{G}\mathbf{C}^{mm} \qquad (3.1.66)$$

$$\mathbf{C}^{md}=\mathbf{C}^{mm}\mathbf{G}^T \qquad (3.1.67)$$

证（3.1.66）：

根据（3.1.56）$\mathbf{C}_0^{dd}=\mathbf{S}-\mathbf{G}\mathbf{C}_0^{mm}\mathbf{G}^T$ 代入（3.1.59）化简后，既得（3.1.66）证毕。

由（3.1.65）－（3.1.67）可知，$\mathbf{C}^{mm}$ 携带了模型评估的所有信息。其对角线表征后验模型 $\tilde{\mathbf{m}}$ 的方差，$\mathbf{A}^{mm}=\mathbf{C}_0^{mm}\mathbf{G}^T\mathbf{S}^{-1}\mathbf{G}$ 表征后验模型的分辨。对于非线性问题，如果观测数据与先验模型参数相互独立，并且数据可用模型参数显式表示，把最后一次得到的 $\mathbf{G}$ 代入到 $\mathbf{C}^{mm}$ 和 $\mathbf{A}^{mm}$ 中，就可以对计算的后验模型进行评估。

3.2 连续问题

3.2.1 连续函数的展开

在离散问题中，要求模型参数是有限维的。在连续问题中需要反演的模型参数是连续函数。通常的做法是把连续函数离散化，然后用离散问题反演方法。比如做全球层析成像，在球面上用球谐函数展开，在深度上用多项式展开。在做一维深度反演时，利用样条函数模拟地震波速度随深度的变化。连续函数的展开需要用到基函数，理论上需要展开到无穷多项，

但在实际问题中是不可能展开到无穷的，这就需要截断，由此引出截断问题。样条函数虽然不存在截断问题，但如果参数用的太少，就不可能很好的模拟一个连续函数。

本节将采用不同的思路，而直接反演连续函数模型。在讨论这个问题时，需要用到希尔伯特空间的概念。先定义函数的内积，设有两个连续的平方可积函数 $f(x_1,\cdots,x_n)$，$g(x_1,\cdots,x_n)$，内积定义为

$$f\cdot g=\int_v f(x_1,\cdots,x_n)g(x_1,\cdots,x_n)dx_1\cdots dx_n \qquad (3.2.1)$$

函数 $f(x_1,\cdots,x_n)$ 的范数定义为

$$\|f\|^2=\int_v f^2(x_1,\cdots,x_n)dx_1\cdots dx_n \qquad (3.2.2)$$

以下为了书写方便，略去积分限，必要的时候再加上。有了长度(范数)和内积的概念，连续平方可积函数就可以看作空间中的矢量。 希尔伯特空间被定义为完备的酉空间。在一定的条件下一个连续函数就是希尔伯特空间中的一个矢量，连续平方可积函数就是满足条件的希尔伯特空间中的矢量。 既然希尔伯特空间是完备的，就一定能够找到一组归一化的相互正交的基函数

$$\left\{\varphi(\mathbf{x})_1,\varphi_2(\mathbf{x}),\cdots\varphi_n(\mathbf{x}),\cdots\right\},$$

$$\varphi_i(\mathbf{x})\cdot\varphi_j(\mathbf{x})=\delta_{ij} \qquad (3.2.3)$$

希尔伯特空间中的任一函数 $f(x_1,\cdots,x_n)$ 都可用这组基来展开

$$f(\mathbf{x})=a_1\varphi_1(\mathbf{x})+a_2\varphi_2(\mathbf{x})+\cdots+a_\alpha\varphi_\alpha(\mathbf{x})+\cdots$$

或表示为 $$f(\mathbf{x})=\sum_{\alpha=1}^{\infty}a_\alpha\varphi_\alpha(\mathbf{x})=a_i\varphi_i(\mathbf{x}) \quad \text{(拉丁字母下脚标求和)} \qquad (3.2.4)$$

根据基函数的正交性可知

$$a_i=f\cdot\varphi_i=\int f(\mathbf{x})\varphi_i(\mathbf{x})dV \qquad (3.2.5)$$

序列 $\{a_i\}$ 称为函数 $f(\mathbf{x})$ 在此基下的坐标。

例如 $$\delta(\mathbf{x}-\mathbf{x}')=\left(\int\delta(\mathbf{x}''-\mathbf{x}')\varphi_i(\mathbf{x}'')dV''\right)\varphi_i(x)=\varphi_i(x)\varphi_i(x')$$

根据线性代数，希尔伯特空间中的任一算子 $B(\mathbf{x})$ 在此基下的矩阵可表示为

$$\beta_{ij}=\varphi_i\cdot\left(B(\mathbf{x})*\varphi_j(\mathbf{x})\right)=\int\varphi_i(\mathbf{x})\cdot\left(B(\mathbf{x})*\varphi_j(\mathbf{x})\right)dV \qquad (3.2.6)$$

算子 $B(\mathbf{x})$ 中的 $\mathbf{x}$ 有两个含义，其一表示算子 B 可能是 $\mathbf{x}$ 的函数，其二表示算子 B 只作用到自变量为 $\mathbf{x}$ 的函数上，算子后面的星号“＊”表示“作用于”，符号“ · ”表示内积。

例如，$B(\mathbf{x})*\left(f(\mathbf{x})g(\mathbf{x}')\right)=g(\mathbf{x}')\,B(\mathbf{x})*f(\mathbf{x})$，算子 $B(\mathbf{x})$ 并不作用到 $g(\mathbf{x}')$ 上。

在函数希尔伯特空间中带参变量 $\mathbf{x}'$ 的 δ（$\mathbf{x}-\mathbf{x}'$）函数的作用相当于单位算子。定义希尔伯特空间中的单位算子 I，

$$I*f(\mathbf{x})=f(\mathbf{x})*I=\int\delta(\mathbf{x}-\mathbf{x}')f(\mathbf{x}')dV'=f(\mathbf{x})$$

$$I*B(\mathbf{x})=B(\mathbf{x})*I=B(\mathbf{x}) \qquad (3.2.7)$$

根据（3.2.6），单位算子 I 在基下的矩阵为

$$\delta_{ij}=\varphi_i(\mathbf{x})\cdot\left(I*\varphi_j(\mathbf{x})\right)=\int\varphi_i(\mathbf{x})\varphi_j(\mathbf{x})dV$$

像定义函数希尔伯特空间中的单位算子一样，我们定义函数希尔伯特空间中的协方差算子 $B(\mathbf{x})$，

$$B(\mathbf{x})*f(\mathbf{x})=\int C(\mathbf{x},\mathbf{x}')f(\mathbf{x}')dV' \qquad (3.2.8)$$

其中 $C(\mathbf{x},\mathbf{x}')$ 为函数 $f(\mathbf{x})$ 的协方差函数。根据定义可知，协方差算子是线性算子。

定义协方差逆算子 $B^{-1}(\mathbf{x})$

$$B(\mathbf{x})*B^{-1}(\mathbf{x})=B^{-1}(\mathbf{x})*B(\mathbf{x})=I \text{ （单位算子）} \qquad (3.2.9)$$

由(3.2.8)和(3.2.9)得

$$B^{-1}(\mathbf{x})*B(\mathbf{x})*f(\mathbf{x})=B^{-1}(\mathbf{x})*\left\{\int C(\mathbf{x},\mathbf{x}')f(\mathbf{x}')dV'\right\}=f(\mathbf{x})$$

令 $$B^{-1}(\mathbf{x})*g(\mathbf{x})=y(\mathbf{x}) \qquad (3.2.10)$$

则 $$g(\mathbf{x})=B(\mathbf{x})*y(\mathbf{x})=\int C(\mathbf{x},\mathbf{x}')y(\mathbf{x}')dV' \qquad (3.2.11)$$

所以，协方差逆算子作用到一个函数上得到的结果要解（3.2.11）的积分方程。

由（3.2.6）得知协方差算子在基 $\{\varphi_i\}$ 下的矩阵为

$$B_{ij}=\varphi_i(\mathbf{x})\cdot\left(B(\mathbf{x})*\varphi_j(\mathbf{x})\right)=\iint\varphi_i(\mathbf{x})C(\mathbf{x},\mathbf{x}')\varphi_j(\mathbf{x}')dV'dV \qquad (3.2.12)$$

协方差逆算子在基下的矩阵为

$$B_{ij}^{-1}=\varphi_i(\mathbf{x})\cdot\left(B^{-1}(\mathbf{x})*\varphi_j(\mathbf{x})\right)=\int\varphi_i(\mathbf{x})B^{-1}(\mathbf{x})*\varphi_j(\mathbf{x})dV$$

如此，任一希尔伯特空间中的函数就有唯一的一个无穷维的数组矢量 $(a_1\ \ a_2\cdots\ a_i\cdots)^T$ 与之对应.。$(a_1\ \ a_2\cdots\ a_i\cdots)^T$ 被称为函数 $f(\mathbf{x})$ 的坐标矢量。 设函数 $f(\mathbf{x})$ 的坐标为 $\mathbf{a}$， $g(\mathbf{x})$ 的坐标为 $\mathbf{b}$。$\mathbf{a}$ 与 $\mathbf{b}$ 的内积定义为

$$\mathbf{a}\cdot\mathbf{b}=\sum_{\beta=1}^{\infty}a_\beta b_\beta=a_i b_i$$

则 $$\mathbf{a}\cdot\mathbf{b}=f(\mathbf{x})\cdot g(\mathbf{x})$$

所以，两个函数的内积就对应于两个函数坐标矢量的内积。 任一希尔伯特空间中的线性算子都有一个矩阵与其相对应。所以，函数的坐标矢量也构成一希尔伯特空间。 两个空间是同构的。 既然如此，那么连续模型函数的反演问题就可以在无穷维的数组空间中讨论。 以下我们将遵从这一思路来讨论连续模型的反演问题。

有了以上的数学准备，我们来描述连续模型参数的正演和反演问题。

设数据是离散的用 d_i $(i=1,\cdots,n)$ 表示，具有先验均值 d_{0i} $(i=1,\cdots,n)$ 和关于 d_{0i} 的先验协方差矩阵 $\mathbf{C}_{d_0}$。连续模型用 $m(\mathbf{x})$ 表示，具有先验均值 $m_0(\mathbf{x})$ 以及关于 $m_0(\mathbf{x})$ 的先验协方差函数 $C_{m_0}(\mathbf{x},\mathbf{x}')$，正演条件为

$$f_i\big(\mathbf{d},m(\mathbf{x})\big)=0\,,\quad(i=1,\cdots,n)\qquad（3.2.13）$$

，在广义最小二乘准则下求模型函数的解 $\tilde{m}(\mathbf{x})$。

设模型空间 $\{\,m(\mathbf{x})\}$ 构成一完备的希尔伯特空间，存在一组正交完备基 $\{\,\varphi(\mathbf{x})_1,\varphi_2(\mathbf{x}),\cdots\varphi_i(\mathbf{x}),\cdots\}$。那么模型函数 $m(\mathbf{x})$ 在此基下就可表为

$$m(\mathbf{x})=p_i\varphi_i(\mathbf{x})\qquad（3.2.14）$$

$$p_i=m(\mathbf{x})\cdot\varphi_i(\mathbf{x})=\int m(\mathbf{x})\varphi_i(\mathbf{x})dV\qquad(3.2.15)$$

先验均值 $m_0(\mathbf{x})$ 可表为

$$m_0(\mathbf{x})=p_{0i}\varphi_i(\mathbf{x})\qquad（3.2.16）$$

$$p_{0i}=m_0(\mathbf{x})\cdot\varphi_i(\mathbf{x})=\int m_0(\mathbf{x})\varphi_i(\mathbf{x})dV\qquad(3.2.17)$$

先验协方差函数对应的先验协方差算子 $B_0(\mathbf{x})$ 在此基下的矩阵为

$$C_{p_0 ij}=\iint\varphi_i(\mathbf{x})C_{m_0}(\mathbf{x},\mathbf{x}')\varphi_j(\mathbf{x}')dV'dV\qquad（3.2.18）$$

则，$\mathbf{C}_{p_0}$ 就是关于 p_{0i} 的先验协方差矩阵。

（习题：证明 $\mathbf{C}_{p_0}$ 是关于 p_{0i} 的协方差矩阵。）

先验协方差逆算子 $B_0^{-1}(\mathbf{x})$ 在此基下的矩阵为

$$C_{p_0 ij}^{-1} = \int \varphi_i(\mathbf{x}) B_0^{-1}(\mathbf{x}) * \varphi_j(\mathbf{x}) dV \qquad (3.2.19)$$

在希尔伯特空间中，应该把 $f(m(\mathbf{x}))$ 看作算子 f 作用到模型 $m(\mathbf{x})$ 上。对 m 的求导

$\dfrac{\partial f(m(\mathbf{x}))}{\partial m}$ 应理解为Freshet导数，$\dfrac{\partial f(m(\mathbf{x}))}{\partial m}$ 也是一个算子，作用到与它有相同自变量的函数上。Freshet导数定义为

$$f(m+\Delta m) - f(m) = \frac{\partial f(m)}{\partial m} * \Delta m$$

例如：$f(m(x)) = \dfrac{dm(x)}{dx}$，根据定义 $\dfrac{d(m+\Delta m)}{dx} - \dfrac{dm}{dx} = \dfrac{d\Delta m}{dx} = \dfrac{d}{dx} * \Delta m$ 。所以Freshet

导数为 $\dfrac{d}{dx}$ 。

将（3.2.14）代入（3.2.13）后，由于函数组 φ_i 是已知的，所以正演条件可等价的表为

$$f_i(d_1, d_2, \cdots, d_n, p_1, p_x, \cdots) = 0 \text{，} (i = 1, \cdots, n) \qquad （3.2.20）$$

令
$$\mathbf{y} = (d_1 \quad d_2 \quad \cdots \quad d_n \quad p_1 \quad p_x \quad \cdots)^T \qquad （3.2.21）$$

先验均值为

$$\mathbf{y}_0 = (d_{01} \quad d_{02} \quad \cdots \quad d_{0n} \quad p_{01} \quad p_{02} \quad \cdots)^T \qquad (3.2.22)$$

则先验协方差矩阵为

$$\mathbf{C}_0 = \begin{pmatrix} \mathbf{C}_{d_0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{p_0} \end{pmatrix} \qquad (3.2.23)$$

此处已假设先验均值中的数据与模型参数不相关。

在广义最小二乘意义下，求模型函数 $\tilde{m}(\mathbf{x})$ 的问题可等价的表为

连续模型离散化后的反演广义最小二乘准则

正问题： $f_i(d_1, d_2, \cdots, d_n, p_1, p_x, \cdots) = 0$，$(i = 1, \cdots, n)$

令 $\mathbf{y} = (d_1 \ d_2 \cdots \ d_n \ \ p_1 \ \ p_x \cdots)^T$（各分量为随机变量序列）

$\mathbf{d} = (d_1 \quad \cdots \quad d_n)^T$，$\mathbf{p} = (p_1 \quad p_2 \quad \cdots)^T$

反问题：已知 $\mathbf{y}$ 的先验均值 $\mathbf{y}_0$（随机变量序列），以及关于 $\mathbf{y}_0$ 的先验协方差

矩阵 $\mathbf{C}_0$，在约束条件 $f_i(\mathbf{y}) = 0$，$(i = 1, \cdots, n)$ 下求满足高斯概率密度

以上表述与离散问题中的表述形式上完全相同，所不同的是随机序列是无穷维的，先验协方差矩阵也是无穷维。因此，它们的解在形式上也应该完全相同。当然，我们并不是真的去解这些无穷维的问题，我们只是想通过这种离散的形式来利用以前我们讨论过的离散结果。最终目的还是要获得连续问题的解。为了彻底解决这个问题，我们还需要几个重要的表达式。

设有函数 $g_i(m)$ 以及 $f_i(\mathbf{d}, m)$，利用（3.2.14）－（3.2.17）求它们的偏微商表达式

$$G_{ij}(\mathbf{p}) = \frac{\partial g_i(m)}{\partial p_j} = \frac{\partial g_i(m(\mathbf{x}))}{\partial m} * \varphi_j(\mathbf{x}) \qquad (3.2.24)$$

$$F_{ij}^{m}(\mathbf{d}, \mathbf{p}) = \frac{\partial f_i(\mathbf{d}, m)}{\partial p_j} = \frac{\partial f_i(\mathbf{d}, m(\mathbf{x}))}{\partial m} * \varphi_j(\mathbf{x})$$

$$F_{ij}^{d}(\mathbf{d}, \mathbf{p}) = \frac{\partial f_i(\mathbf{d}, m(\mathbf{x}))}{\partial d_j} * I \qquad (3.2.25)$$

$$G_{ij}(\mathbf{p})\left(p_j - p_{0j}\right) = \frac{\partial g_i(m(\mathbf{x}))}{\partial m} * \varphi_j(\mathbf{x})\left(\int m(\mathbf{x}'')\varphi_j(\mathbf{x}'')dV'' - \int m_0(\mathbf{x}')\varphi_j(\mathbf{x}')dV'\right)$$

$$= \frac{\partial g_i(m(\mathbf{x}))}{\partial m} * \left(\int m(\mathbf{x}'')\varphi_j(\mathbf{x})\varphi_j(\mathbf{x}'')dV'' - \int m_0(\mathbf{x}')\varphi_j(\mathbf{x})\varphi_j(\mathbf{x}')dV'\right)$$

$$= \frac{\partial g_i(m(\mathbf{x}))}{\partial m} * \left(m(\mathbf{x}) - m_0(\mathbf{x})\right) \qquad (3.2.26)$$

首先讨论连续模型在广义最小二乘准则下反演问题是怎么表述的。根据 （3.2.21）－(3.2.23)

$$(\mathbf{y} - \mathbf{y}_0)^T \mathbf{C}_0^{-1} (\mathbf{y} - \mathbf{y}_0)$$

$$= \left((\mathbf{d} - \mathbf{d_0})^T \quad (\mathbf{p} - \mathbf{p}_0)^T\right) \begin{pmatrix} \mathbf{C}_{d_0}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{p_0}^{-1} \end{pmatrix} \left((\mathbf{d} - \mathbf{d_0}) \quad (\mathbf{p} - \mathbf{p}_0)\right)$$

$$= (\mathbf{d}-\mathbf{d_0})^T \mathbf{C}_{d_0}^{-1}(\mathbf{d}-\mathbf{d_0}) + (\mathbf{p}-\mathbf{p}_0)^T \mathbf{C}_{p_0}^{-1}(\mathbf{p}-\mathbf{p}_0) \qquad (3.2.27)$$

下脚标形式为

$$(y_i - y_{0i})(C_{p_0ij}^{-1})(y_j - y_j) = (d_i - d_{0i})(C_{d_0ij}^{-1})(d_j - d_{0j}) + (p_i - p_{0i})C_{p_0ij}^{-1}(p_j - p_{0j})$$

又根据（3.2.16）－（3.2.19）

$$(p_i - p_{0i})C_{p_0ij}^{-1}(p_j - p_{0j})$$

$$= \int \big(m(\mathbf{x}') - m_0(\mathbf{x}')\big)\varphi_i(\mathbf{x}')dV' \int \varphi_i(\mathbf{x})B_0^{-1}(\mathbf{x}) * \varphi_j(\mathbf{x})dV \int \varphi_j(\mathbf{x}'')\big(m(\mathbf{x}'') - m_0(\mathbf{x}'')dV''\big)$$

$$= \iiint \big(m(\mathbf{x}') - m_0(\mathbf{x}')\big)\varphi_i(\mathbf{x}')\varphi_i(\mathbf{x})B_0^{-1}(\mathbf{x}) * \varphi_j(\mathbf{x})\varphi_j(\mathbf{x}'')\big(m(\mathbf{x}'') - m_0(\mathbf{x}'')\big)dVdV'dV''$$

$$= \iint \big(m(\mathbf{x}) - m_0(\mathbf{x})\big)B_0^{-1}(\mathbf{x}) * \varphi_j(\mathbf{x})\varphi_j(\mathbf{x}'')\big(m(\mathbf{x}'') - m_0(\mathbf{x}'')\big)dV''dV$$

$$= \int \big(m(\mathbf{x}) - m_0(\mathbf{x})\big)B_0^{-1}(\mathbf{x}) * \Big(\int \varphi_j(\mathbf{x})\varphi_j(\mathbf{x}'')\big(m(\mathbf{x}'') - m_0(\mathbf{x}'')\big)dV''\Big)dV$$

$$= \int \big(m(\mathbf{x}) - m_0(\mathbf{x})\big)B_0^{-1}(\mathbf{x}) * \big(m(\mathbf{x}) - m_0(\mathbf{x})\big)dV$$

$$= \big(m(\mathbf{x}) - m_0(\mathbf{x})\big)\cdot B_0^{-1}(\mathbf{x}) * \big(m(\mathbf{x}) - m_0(\mathbf{x})\big) \qquad (3.2.28)$$

利用（3.2.27）,(3.2.28)就可以将连续模型在广义最小二乘准则下反演的问题表述如下

连续模型反演广义最小二乘准则

正问题：　$f_i\big(\mathbf{d}, m(\mathbf{x})\big) = 0$，$(i = 1, \cdots, n)$

令　$\mathbf{d} = \begin{pmatrix} d_1 & d_2 & \cdots & d_n \end{pmatrix}^T$（随机变量序列）。

反问题：已知$\mathbf{d}$的先验均值$\mathbf{d}_0$（随机变量序列），以及关于$\mathbf{d}_0$的先验协方差矩阵

$\mathbf{C}_{d_0}$，在约束条件$f_i\big(\mathbf{d}, m(\mathbf{x})\big) = 0$，$(i = 1, \cdots, n)$下求满足高斯概率密度

$$\rho(\mathbf{d}, \mathbf{m}) = const \cdot \exp\left\{-\frac{1}{2}\Big((\mathbf{d}-\mathbf{d}_0)^T \mathbf{C}_0^{-1}(\mathbf{d}-\mathbf{d}_0) + (m - m_0)\cdot B_0^{-1} * (m - m_0)\Big)\right\}$$

极大值的解$\mathbf{d} = \tilde{\mathbf{d}}$和$m(\mathbf{x}) = \tilde{m}(\mathbf{x})$。

连续问题通过基函数展开离散化后就可以利用离散问题的结果得到其解。

根据（3.1.31）－（3.1.34）以及“连续模型离散化后的反演广义最小二乘准则”表述得

$$F_{ij}^d = \frac{\partial f_i(\mathbf{y})}{\partial d_j}, \quad F_{ij}^m = \frac{\partial f_i(\mathbf{y})}{\partial p_j} \qquad (3.2.29)$$

$$\mathbf{S}(\mathbf{y})=\mathbf{F}^d(\mathbf{y})\mathbf{C}_{d_0}\left(\mathbf{F}^d(\mathbf{y})\right)^T+\mathbf{F}^m(\mathbf{y})\mathbf{C}_{p_0}\left(\mathbf{F}^m(\mathbf{y})\right)^T \quad (3.2.30)$$

$$\begin{cases}\mathbf{d}^{(n+1)}=\mathbf{d}_0+\mathbf{C}_{d_0}\left(\mathbf{F}^d(\mathbf{y}^{(n)})\right)^T\mathbf{S}^{-1}(\mathbf{y}^{(n)})\left\{\left(\mathbf{F}^d(\mathbf{y}^{(n)}) \quad \mathbf{F}^m(\mathbf{y}^{(n)})\right)\begin{pmatrix}\mathbf{d}-\mathbf{d}_0\\ \mathbf{p}-\mathbf{p}_0\end{pmatrix}-\mathbf{f}(\mathbf{y}^{(n)})\right\}\\ \mathbf{p}^{(n+1)}=\mathbf{p}_0+\mathbf{C}_{p_0}\left(\mathbf{F}^m(\mathbf{y}^{(n)})\right)^T\mathbf{S}^{-1}(\mathbf{y}^{(n)})\left\{\left(\mathbf{F}^d(\mathbf{y}^{(n)}) \quad \mathbf{F}^m(\mathbf{y}^{(n)})\right)\begin{pmatrix}\mathbf{d}-\mathbf{d}_0\\ \mathbf{p}-\mathbf{p}_0\end{pmatrix}-\mathbf{f}(\mathbf{y}^{(n)})\right\}\end{cases}$$

（3.2.31）

对线性问题

正问题：$\mathbf{f}(\mathbf{d},\mathbf{p})=\mathbf{F}\begin{pmatrix}\mathbf{d}\\ \mathbf{p}\end{pmatrix}=\mathbf{0}$，算子 $\mathbf{F}$ 不含 $\mathbf{d}, \mathbf{p}$ 的分量。

反问题解：

$$\mathbf{S}=\mathbf{F}^d\mathbf{C}_{d_0}\left(\mathbf{F}^d\right)^T+\mathbf{F}^m\mathbf{C}_{p_0}\left(\mathbf{F}^m\right)^T$$

$$\tilde{\mathbf{d}}=\mathbf{d}_0-\mathbf{C}_{d_0}\left(\mathbf{F}^d\right)^T\mathbf{S}^{-1}\left(\mathbf{F}^d\mathbf{d}_0+\mathbf{F}^m\mathbf{p}_0\right) \quad (3.2.32)$$

$$\tilde{\mathbf{p}}=\mathbf{p}_0-\mathbf{C}_{p_0}\left(\mathbf{F}^m\right)^T\mathbf{S}^{-1}\left(\mathbf{F}^d\mathbf{d}_0+\mathbf{F}^m\mathbf{p}_0\right) \quad (3.2.33)$$

根据（3.2.14）－（3.2.19）以及（3.2.24）－（3.2.26），代入（3.2.29）－（3.2.33）就得到连续问题的解。以下分几种情况讨论。

1) 一般情况

用下脚表表示

$$F_{ij}^d(\mathbf{d},\mathbf{p})=\frac{\partial f_i(\mathbf{d},m)}{\partial d_j}*I,\quad F_{ij}^m(\mathbf{d},\mathbf{p})=\frac{\partial f_i(\mathbf{d},m(\mathbf{x}))}{\partial p_j}=\frac{\partial f_i(\mathbf{d},m(\mathbf{x}))}{\partial m}*\varphi_j(\mathbf{x})$$

（3.2.34）

$$S_{ij}=F_{ik}^d(\mathbf{d},\mathbf{p})C_{d_0kl}F_{jl}^d(\mathbf{d},\mathbf{p})+F_{ik}^m(\mathbf{d},\mathbf{p})C_{p_0kl}F_{jl}^m(\mathbf{d},\mathbf{p})$$

$$=F_{ik}^d(\mathbf{d},m(\mathbf{x}'))C_{d_0kl}F_{jl}^d(\mathbf{d},m(\mathbf{x}''))$$

$$+\frac{\partial f_i(\mathbf{d},m(\mathbf{x}))}{\partial m}*\varphi_k(\mathbf{x})\iint\varphi_k(\mathbf{x}'')C_{m_0}(\mathbf{x}'',\mathbf{x}''')\varphi_l(\mathbf{x}''')dV''dV'''\frac{\partial f_j(\mathbf{d},m(\mathbf{x}'))}{\partial m}*\varphi_l(\mathbf{x}')$$

$$=F_{ik}^d(\mathbf{d},m(\mathbf{x}'))C_{d_0kl}F_{jl}^d(\mathbf{d},m(\mathbf{x}''))+\frac{\partial f_i(\mathbf{d},m(\mathbf{x}'))}{\partial m}*C_{m_0}(\mathbf{x}',\mathbf{x}'')*\frac{\partial f_j(\mathbf{d},m(\mathbf{x}''))}{\partial m} \quad (3.2.35)$$

$$d_i^{(n+1)}=d_{0i}+C_{d_0ij}F_{kj}^d(\mathbf{d}^{(n)},m^{(n)}(\mathbf{x}))S_{kl}^{-1}(\mathbf{d}^{(n)},m^{(n)}(\mathbf{x}))$$

$$\left\{F_{lm}^d(\mathbf{d}^{(n)},m^{(n)}(\mathbf{x}))(d_m^{(n)}-d_{0m})+\frac{\partial f_l(\mathbf{d},m^{(n)}(\mathbf{x}'))}{\partial m}*(m^{(n)}(\mathbf{x}')-m_0(\mathbf{x}'))-f_l(\mathbf{d}^{(n)},m^{(n)}(\mathbf{x}))\right\}$$

(3.2.36)

$$p_i^{(n+1)} = p_{0i} + C_{p_0ij} F_{kj}^m(\mathbf{d}^{(n)}, \mathbf{p}^{(n)})) S_{kl}^{-1}(\mathbf{d}^{(n)}, m^{(n)})$$

$$\left\{ F_{lm}^d(\mathbf{d}^{(n)}, m^{(n)}(\mathbf{x}'))(d_m^{(n)} - d_{0m}) + \frac{\partial f_l(\mathbf{d}, m^{(n)}(\mathbf{x}'))}{\partial m} * (m^{(n)}(\mathbf{x}') - m_0(\mathbf{x}')) - f_l(\mathbf{d}^{(n)}, m^{(n)}(\mathbf{x})) \right\}$$

(3.2.37)

用$\varphi_i(\mathbf{x})$乘（3.2.37）再利用（3.2.14）－（3.2.19）得

$$m^{(n+1)}(\mathbf{x}) = m_0(\mathbf{x}) + \left(\frac{\partial f_i(\mathbf{d}\ , m^{(n)}(\mathbf{x}'))}{\partial m} * C_{m_0}(\mathbf{x}, \mathbf{x}') \right) S_{ij}^{-1}(\mathbf{d}^{(n)}, m^{(n)})$$

$$\left\{ F_{jk}^d(\mathbf{d}^{(n)}, m^{(n)}(\mathbf{x}'))(d_k^{(n)} - d_{0k}) + \frac{\partial f_j(\mathbf{d}, m^{(n)}(\mathbf{x}'))}{\partial m} * (m^{(n)}(\mathbf{x}') - m_0(\mathbf{x}')) - f_j(\mathbf{d}^{(n)}, m^{(n)}(\mathbf{x})) \right\}$$

（3.2.38）

对于非线性问题，（3.2.36）和（3.2.38）是迭代过程。如果收敛则$\left\| \mathbf{d}^{(k)} - \mathbf{d}_0 \right\|$将趋于平稳。

对线性问题

正问题：$f_i(\mathbf{d}, \mathbf{p}) = (\mathbf{F}^d \quad \mathbf{F}^m) \begin{pmatrix} \mathbf{d} \\ \mathbf{p} \end{pmatrix} = \mathbf{0}$

即 $F_{ij}^d d_j + F_{ij}^m p_j = 0$ （3.2.39）

算子F_{ij}^d和F_{ij}^m不含$\mathbf{d}$和$\mathbf{p}$的分量。

利用（3.2.34），一般的线性系统（3.2.39）可表为

$$F_{ij}^d d_j + \frac{\partial f_i(\mathbf{d}\ , m(\mathbf{x}))}{\partial m} * m(\mathbf{x}) = 0$$

或 $F_{ij}^d d_j + g_i(\mathbf{x}) * m(\mathbf{x}) = 0$ （3.2.40）

其中算子$g_i(\mathbf{x}) = \dfrac{\partial f_i(\mathbf{d}\ , m(\mathbf{x}))}{\partial m}$不含模型函数$m(\mathbf{x})$和$\mathbf{d}$的分量。

反问题解：

$$S_{ij} = F_{ik}^d C_{d_0kl} F_{jl}^d + g_i(\mathbf{x}) * C_{m_0}(\mathbf{x}, \mathbf{x}') * g_j(\mathbf{x}') \qquad （3.2.41）$$

$$\tilde{d}_i = d_{0i} - C_{d_0ij} F_{kj}^d S_{kl}^{-1} \left\{ F_{lm}^d d_{0m} + g_l(\mathbf{x}') * m_0(\mathbf{x}') \right\} \qquad （3.2.42）$$

$$\tilde{m}(\mathbf{x}) = m_0(\mathbf{x}) - \left(g_i(\mathbf{x}') * C_{m_0}(\mathbf{x}, \mathbf{x}') \right) S_{ij}^{-1} \left\{ F_{jk}^d d_{0k} + g_j(\mathbf{x}') * m_0(\mathbf{x}') \right\} \qquad (3.2.43)$$

线性问题无需迭代。

2) 数据可以用模型函数显式表示

根据（3.1.35）—（3.1.41）可知

正问题：$d_i - g_i(\mathbf{p}) = 0$，$(i=1,\cdots,n)$

变换后

$$d_i - g_i(m(\mathbf{x})) = 0,\ (i=1,\cdots,n) \qquad (3.2.44)$$

反问题解：

$$F_{ij}^d = \delta_{ij},\ F_{ij}^m(\mathbf{p}) = -G_{ij}(\mathbf{p}),\ G_{ij}(\mathbf{p}) = \frac{\partial g_i(m(\mathbf{x}))}{\partial m} * \varphi_j(\mathbf{x}) \qquad (3.2.45)$$

$$S_{ij} = C_{d_0 ij} +$$

$$\frac{\partial g_i(m(\mathbf{x}))}{\partial m} * \varphi_l(\mathbf{x}) \iint \varphi_l(\mathbf{x}'') C_{m_0}(\mathbf{x}'', \mathbf{x}''') \varphi_k(\mathbf{x}''') dV'' dV''' \frac{\partial g_j(m(\mathbf{x}'))}{\partial m} * \varphi_k(\mathbf{x}')$$

$$= C_{d_0 ij} + \frac{\partial g_i(m(\mathbf{x}))}{\partial m} * C_{m_0}(\mathbf{x}, \mathbf{x}') * \frac{\partial g_j(m(\mathbf{x}'))}{\partial m} \qquad (3.2.46)$$

$$d_i^{(n+1)} = d_{0i}$$

$$- C_{d_0 ij} S_{jk}^{-1}(m^{(n)}) \left\{ (d_{0k} - d_k^{(n)}) + \frac{\partial g_k(m^{(n)}(\mathbf{x}'))}{\partial m} * (m^{(n)}(\mathbf{x}') - m_0(\mathbf{x}')) \right\} \qquad (3.2.47)$$

$$m^{(n+1)}(\mathbf{x}) = m_0(\mathbf{x}) + \left(C_{m_0}(\mathbf{x}, \mathbf{x}') * \frac{\partial g_j(m^{(n)}(\mathbf{x}'))}{\partial m} \right)$$

$$S_{jk}^{-1}(m^{(n)}) \left\{ (d_{0k} - d_k^{(n)}) + \frac{\partial g_k(m^{(n)}(\mathbf{x}'))}{\partial m} * (m^{(n)}(\mathbf{x}') - m_0(\mathbf{x}')) \right\} \qquad (3.2.48)$$

对于线性问题

$$g_i(m(\mathbf{x})) = g_i(\mathbf{x}) * m(\mathbf{x})$$

正问题：$d_i = g_i(\mathbf{x}) * m(\mathbf{x})$ （3.2.49）

其中算子 $g_i(\mathbf{x})$ 不含模型函数和数据分量。

反问题解：

由（3.2.45）得

$$\frac{\partial g_i(m(\mathbf{x}))}{\partial m} = g_i(\mathbf{x})$$

代入（3.2.46）－（3.2.48）得

$$S_{ij}=C_{d_0ij}+g_i(\mathbf{x}')*C_{m_0}(\mathbf{x}',\mathbf{x}'')*g_j(\mathbf{x}'') \qquad (3.2.50)$$

$$\tilde{d}_i=d_{0i}-C_{d_0ij}S_{jk}^{-1}\{\, d_{0k}-g_k(\mathbf{x}')*m_0(\mathbf{x}')\,\}=g_i(\mathbf{x}')*\tilde{m}(\mathbf{x}') \qquad (3.2.51)$$

$$\tilde{m}(\mathbf{x})=m_0(\mathbf{x})+\left(g_j(\mathbf{x}')*C_{m_0}(\mathbf{x},\mathbf{x}')\right)S_{jk}^{-1}\{\, d_{0k}-g_k(\mathbf{x}')*m_0(\mathbf{x}')\,\} \qquad (3.2.52)$$

3) 无数据

正问题：

$$f_i(m(\mathbf{x}))=0\,,\quad (i=1,\cdots n) \qquad （3.2.53）$$

根据 (3.2.35)（3.2.38）可知

$$S_{ij}=\frac{\partial f_i(\,m(\mathbf{x}'))}{\partial m}*C_{m_0}(\mathbf{x}',\mathbf{x}'')*\frac{\partial f_j(\,m(\mathbf{x}''))}{\partial m} \qquad (3.2.54)$$

$$m^{(n+1)}(\mathbf{x})=m_0(\mathbf{x})+\left(\frac{\partial f_i(\,m^{(n)}(\mathbf{x}'))}{\partial m}*C_{m_0}(\mathbf{x},\mathbf{x}')\right)S_{ij}^{-1}(\,m^{(n)})$$
$$\left\{\frac{\partial f_j(\,m^{(n)}(\mathbf{x}'))}{\partial m}*(m^{(n)}(\mathbf{x}')-m_0(\mathbf{x}'))-f_j(\,m^{(n)}(\mathbf{x}))\right\} \qquad (3.2.55)$$

4) 分辨和误差分析

如果迭代收敛，收敛前的几次迭代结果应该非常接近，就可以做线性化近似。所以最后模型的评估就可以用线性系统中的评估方法。

设数据可以用模型函数显式表示，正问题为：$d_i-g_i(m(\mathbf{x}))=0$

根据（3.1.58）模型的后验协方差矩阵$\mathbf{C}_p$为

$$C_{pij}=C_{p_0ij}-C_{p_0ik}G_{lk}S_{lm}^{-1}G_{mn}C_{p_0nj} \qquad （3.2.56）$$

设后验协方差函数为$C^m(\mathbf{x},\mathbf{x}')$，根据协方差算子的定义（3.2.8）和（3.2.12）

$$\iint\varphi_i(\mathbf{x}'')C^m(\mathbf{x}'',\mathbf{x}''')\varphi_j(\mathbf{x}''')dV''dV'''=\iint\varphi_i(\mathbf{x}'')C_{m_0}(\mathbf{x}'',\mathbf{x}''')\varphi_j(\mathbf{x}''')dV''dV'''$$
$$-\iint\varphi_i(\mathbf{x}'')C_{m_0}(\mathbf{x}'',\mathbf{x}''')\varphi_k(\mathbf{x}''')dV''dV'''\frac{\partial g_l(m(\mathbf{x}))}{\partial m}*\varphi_k(\mathbf{x})S_{lm}^{-1}$$
$$\frac{\partial g_m(m(\mathbf{x}))}{\partial m}*\varphi_n(\mathbf{x})\iint\varphi_n(\mathbf{x}'')C_{m_0}(\mathbf{x}'',\mathbf{x}''')\varphi_j(\mathbf{x}''')dV''dV'''$$

上式左乘$\varphi_i(\mathbf{x})$，右乘$\varphi_j(\mathbf{x}')$，并利用$\delta(\mathbf{x}-\mathbf{x}')=\varphi_i(\mathbf{x})\varphi_i(\mathbf{x}')$得

$$C^m(\mathbf{x},\mathbf{x}')=C_{m_0}(\mathbf{x},\mathbf{x}')-C_{m_0}(\mathbf{x},\mathbf{x}'')*\frac{\partial g_l(m(\mathbf{x}''))}{\partial m}S_{lm}^{-1}\frac{\partial g_m(m(\mathbf{x}''))}{\partial m}*C_{m_0}(\mathbf{x}'',\mathbf{x}')$$

$$= C_{m_0}(\mathbf{x}, \mathbf{x}') - A^m(\mathbf{x}, \mathbf{x}') \qquad (3.2.57)$$

其中

$$A^m(\mathbf{x}, \mathbf{x}') = C_{m_0}(\mathbf{x}, \mathbf{x}'') * \frac{\partial g_l(m(\mathbf{x}''))}{\partial m} S_{lm}^{-1} \frac{\partial g_m(m(\mathbf{x}''))}{\partial m} * C_{m_0}(\mathbf{x}'', \mathbf{x}') \qquad (3.2.58)$$

$$S_{ij} = C_{d_0 ij} + \frac{\partial g_i(m(\mathbf{x}'))}{\partial m} * C_{m_0}(\mathbf{x}', \mathbf{x}'') * \frac{\partial g_j(m(\mathbf{x}''))}{\partial m} \qquad (3.2.59)$$

(3.2.57) －（3.2.59）右边各量用（3.2.46）－（3.2.48）最后一次迭代的值。

后验协方差函数既是分辨的度量又是后验模型方差的度量，当 $\mathbf{x} = \mathbf{x}'$ 时，它表示该点的方差，当 $\mathbf{x} \neq \mathbf{x}'$ 时，它为模型的分辨。计算结果通常像图7。

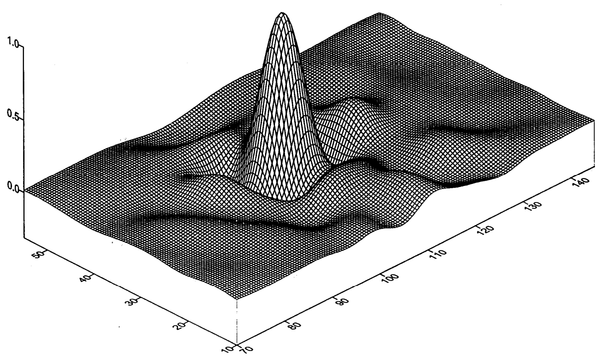

图 7 模型的分辨图

3.3 举例

1) 无网格走时反演

正演条件为：$t_i = \int\limits_{s(v)} \frac{1}{v(\mathbf{x})} ds_i$，$(i = 1, \cdots, n)$ (3.3.1)

其中 $v(\mathbf{x})$ 为速度函数。积分限 $s(v)$ 表示积分路径与介质的速度有关。

写成算子的形式为

$$t_i = \int\limits_{s(m)} ds_i * m(\mathbf{x}),\ (i = 1, \cdots, n), \qquad (3.3.2)$$

其中 $m(\mathbf{x}) = \dfrac{1}{v(\mathbf{x})}$ 为波幔度函数。

反问题：有一组观测数据 t_i^0，$(i = 1, \cdots n)$，其先验协方差矩阵为 $C_{ij}^{d_0}$。先验模型函数为

$m_0(\mathbf{x})$，先验协方差函数为 $C^{m_0}(\mathbf{x},\mathbf{x}')$，求波慢速函数 $\widetilde{m}(\mathbf{x})$ 。

这是一个模型函数，数据可用模型函数显式表示的连续函数反演问题，目标函数为波幔度 $m(\mathbf{x})=\dfrac{1}{v(\mathbf{x})}$。由于算子 $\int\limits_{s(m)} ds_i$ 含有模型函数，所以该问题属于非线性反演。反演公式应该用(3.2.45)－(3.2.48)。

先求Freshet导数。令 $g_i(m(\mathbf{x}))=\int\limits_{s(v)} ds_i * m(\mathbf{x})$，根据Freshet导数定义

$$\int\limits_{s(m+\Delta m)} ds_i *(m(\mathbf{x})+\Delta m)-\int\limits_{s(m)} ds_i * m(\mathbf{x})$$

$$=\int\limits_{s(m)} ds_i *(m(\mathbf{x})+\Delta m)-\int\limits_{s(m)} ds_i * m(\mathbf{x})+o(\Delta m^2) \quad \text{（应用费尔马原理）}$$

$$=\int\limits_{s(m)} ds_i * \Delta m$$

所以得

$$\frac{\partial g_i(m(\mathbf{x}))}{\partial m}=\int\limits_{s(m)} ds_i \tag{3.3.3}$$

根据(3.2.46)得

$$S_{ij}=C_{ij}^{d_0}+\int\limits_{s''(m)}\int\limits_{s'(m)} C^{m_0}(\mathbf{x}',\mathbf{x}'')ds_i' ds_j''$$

根据(3.2.47)得

$$t_i^{(n+1)}=t_i^0-C_{ij}^{d_0}S_{jk}^{-1}(m^{(n)})\left\{(t_k^0-t_k^{(n)})+\int\limits_{s'(m^{(n)})}(m^{(n)}(\mathbf{x}')-m_0(\mathbf{x}'))ds_k'\right\}$$

$$=t_i^0-C_{ij}^{d_0}S_{jk}^{-1}(m^{(n)})\left\{t_k^0-\int\limits_{s'(m^{(n)})} m_0(\mathbf{x}')ds_k'\right\}$$

根据(3.2.48)得

$$m^{(n+1)}(\mathbf{x})=m_0(\mathbf{x})+\left(\int\limits_{s'(m^{(n)})} C^{m_0}(\mathbf{x},\mathbf{x}')ds_j'\right)S_{jk}^{-1}(m^{(n)})\left\{t_k^0-\int\limits_{s'(m^{(n)})} m_0(\mathbf{x}')ds_k'\right\}$$

根据(3.2.57)得模型函数得后验协方差

$$C^m(\mathbf{x},\mathbf{x}')=C^{m_0}(\mathbf{x},\mathbf{x}')-\int\limits_{s''(m)} C^{m_0}(\mathbf{x},\mathbf{x}'')ds_i'' S_{ij}^{-1}\int\limits_{s''(m)} C^{m_0}(\mathbf{x}'',\mathbf{x}')ds_j''$$

以上的积分都是对射线路径的积分。

习题：当目标函数为速度 $v(\mathbf{x})$ 时，写出反演公式及后验协方差函数。

费尔马原理

给定震源，接受点以及地震波速度。定义地震波走时

$$T=\int_{s(v)}\frac{1}{v(\mathbf{x})}ds$$

，走时 T 是射线路径 s 的泛函。

费尔马原理：物理上可行的路径是使泛函 T 取极值的路径。

泛函极值的必要条件是：泛函 T 的变分为零。

2) 面波群速度反演

假设面波行走的路径为地球上的大圆弧。

正演条件为：$t_i=\int_s\frac{1}{v(\mathbf{x})}ds_i$，$(i=1,\cdots,n)$

其中 $v(\mathbf{x})$ 为面波群速度，t_i 为第 i 条路径的群速度走时。积分路径为震中点到地震台站之间的大圆弧。

反问题：已知某一区域一组观测的面波群速度走时数据 t_i^0，$(i=1,\cdots,n)$、观测数据的协方差矩阵 $C_{ij}^{t_0}$，以及先验群速度 $v_0(\mathbf{x})$ 和关于先验群速度的协方差函数 $C^{v_0}(\mathbf{x},\mathbf{x}')$。求目标函数 $v(\mathbf{x})$。

3.4 多个模型函数的情况

以上讨论的是一个模型函数的情况。在实际问题中会碰到一个物理模型用多个连续函数来描述的情况。比如反演速度随深度的变化，一般是把深度按间断面分层，每层可以用一个速度函数来描述。又比如，我们不仅要反演地震P波速度，同时还要反演S波速度和密度。这样就可能需要三个模型函数。先讨论两个模型函数的情况，然后再推广到多个。

设模型函数为 $m_1(\mathbf{x})$ 和 $m_2(\mathbf{x})$。它们属于同一个希尔伯特空间，具有先验均值 $m_1^0(\mathbf{x})$，$m_2^0(\mathbf{x})$，以及关于 $m_1^0(\mathbf{x})$ 和 $m_2^0(\mathbf{x})$ 的先验协方差函数 $C_1^{m_0}(\mathbf{x},\mathbf{x}')$ 和 $C_2^{m_0}(\mathbf{x},\mathbf{x}')$。为简单起

见，设两个先验模型函数之间不相关。数据是离散的，d_i，$(i=1,\cdots,n)$， 先验均值为 d_i^0 。先验协方差矩阵为 $C_{ij}^{d_0}$ 。设先验均值 d_i^0 与先验模型函数之间不相关。

正演条件为

$$f_i(\mathbf{d}, m_1(\mathbf{x}), m_2(\mathbf{x}))=0 ，(i=1,\cdots,n) \qquad （3.4.1）$$

设 $m_1(\mathbf{x})$ 和 $m_2(\mathbf{x})$ 在基函数下展开为

$$m_1(\mathbf{x})=p_i^1\varphi_i(\mathbf{x})， \quad m_2(\mathbf{x})=p_i^2\varphi_i(\mathbf{x}) \qquad （3.4.2）$$

$$m_1^0(\mathbf{x})=p_{0i}^1\varphi_i(\mathbf{x})， \quad m_{2\,02}^0(\mathbf{x})=p_{0i}^2\varphi_i(\mathbf{x}) \qquad （3.4.3）$$

则正演条件可表为

$$f_i(d_1,d_2,\cdots d_n, \quad p_1^1,p_2^1\cdots \quad ,p_1^2,p_2^2\cdots)=0 \qquad (i=1,\cdots,n) \quad （3.4.4）$$

关于 p_{0i}^1 和 p_{0i}^2 的协方差矩阵为

$$C_{1ij}^{p_0}=\iint\varphi_i(\mathbf{x})C_1^{m_0}(\mathbf{x},\mathbf{x}')\varphi_j(\mathbf{x}')dVdV'$$

$$C_{2ij}^{p_0}=\iint\varphi_i(\mathbf{x})C_2^{m_0}(\mathbf{x},\mathbf{x}')\varphi_j(\mathbf{x}')dVdV'$$

则数据和模型分量的协方差矩阵分别为 $C_{ij}^{d_0}$， $C_{1ij}^{p_0}$， $C_{2ij}^{p_0}$，总的先验协方差矩阵为

$$\mathbf{C}_0=\begin{pmatrix}\mathbf{C}^{d_0} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{C}_1^{p_0} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{C}_2^{p_0}\end{pmatrix} \qquad （3.4.5）$$

令 $$\mathbf{b}=(d_1\ d_2\ \cdots d_n\ p_1^1\ p_2^1\cdots p_1^2\ \ p_2^2\cdots)^T \qquad （3.4.6）$$

$$\mathbf{b}_0=(d_1^0\ d_2^0\ \cdots d_n^0\ p_{01}^1\ \ p_{02}^1\cdots p_{01}^2\ \ p_{02}^2\cdots)^T \qquad （3.4.7）$$

$$\mathbf{S}(\mathbf{b})=\mathbf{F}(\mathbf{b})\mathbf{C}_0\mathbf{F}^T(\mathbf{b})$$

$$=\left(\mathbf{F}^d(\mathbf{b})\quad \mathbf{F}^{m_1}(\mathbf{b})\quad \mathbf{F}^{m_2}(\mathbf{b})\right)\begin{pmatrix}\mathbf{C}^{d_0} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{C}_1^{p_0} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{C}_2^{p_0}\end{pmatrix}\begin{pmatrix}\left(\mathbf{F}^d(\mathbf{b})\right)^T\\ \left(\mathbf{F}^{m_1}(\mathbf{b})\right)^T\\ \left(\mathbf{F}^{m_2}(\mathbf{b})\right)^T\end{pmatrix}$$

$$=\mathbf{F}^d(\mathbf{b})\mathbf{C}^{d_0}\left(\mathbf{F}^d(\mathbf{b})\right)^T+\mathbf{F}^{m_1}(\mathbf{b})\mathbf{C}_1^{p_0}\left(\mathbf{F}^{m_1}(\mathbf{b})\right)^T+\mathbf{F}^{m_2}(\mathbf{b})\mathbf{C}_2^{p_0}\left(\mathbf{F}^{m_2}(\mathbf{b})\right)^T \quad （3.4.8）$$

$$F_{ij}(\mathbf{b})=\frac{\partial f_i(\mathbf{b})}{\partial b_j}， \quad F_{ij}^{m_1}(\mathbf{b})=\frac{\partial f_i(\mathbf{b})}{\partial p_j^1}=\frac{\partial f_i(\mathbf{d},m_1,m_2)}{\partial m_1}*\varphi_j(\mathbf{x})，（3.4.9）$$

$$F_{ij}^{d}(\mathbf{b})=\frac{\partial f_i(\mathbf{b})}{\partial d_j}*I\text{，}\quad F_{ij}^{m_2}(\mathbf{b})=\frac{\partial f_i(\mathbf{b})}{\partial p_j^2}=\frac{\partial f_i(\mathbf{d},m_1,m_2)}{\partial m_2}*\varphi_j(\mathbf{x})$$

（3.4.10）

根据（3.1.20）得

$$\mathbf{b}=\mathbf{b}_0+\mathbf{C}_0\mathbf{F}^T(\mathbf{b})\left(\mathbf{F}(\mathbf{b})\mathbf{C}_0\mathbf{F}^T(\mathbf{b})\right)^{-1}\left\{\mathbf{F}(\mathbf{b})(\mathbf{b}-\mathbf{b}_0)-\mathbf{f}(\mathbf{b})\right\}$$

或

$$\begin{pmatrix}\mathbf{d}\\ \mathbf{p}^1\\ \mathbf{p}^2\end{pmatrix}=\begin{pmatrix}\mathbf{d}^0\\ \mathbf{p}_0^1\\ \mathbf{p}_0^2\end{pmatrix}+\begin{pmatrix}\mathbf{C}^{d_0} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{C}_1^{p_0} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{C}_2^{p_0}\end{pmatrix}\begin{pmatrix}\left(\mathbf{F}^d(\mathbf{b})\right)^T\\ \left(\mathbf{F}^{m_1}(\mathbf{b})\right)^T\\ \left(\mathbf{F}^{m_2}(\mathbf{b})\right)^T\end{pmatrix}$$

$$\mathbf{S}^{-1}\left\{\begin{pmatrix}\mathbf{F}^d(\mathbf{b}) & \mathbf{F}^{m_1}(\mathbf{b}) & \mathbf{F}^{m_2}(\mathbf{b})\end{pmatrix}\begin{pmatrix}\mathbf{d}-\mathbf{d}^0\\ \mathbf{p}^1-\mathbf{p}_0^1\\ \mathbf{p}^2-\mathbf{p}_0^2\end{pmatrix}-\mathbf{f}(\mathbf{b})\right\}$$

所以

$$\mathbf{d}=\mathbf{d}^0+\mathbf{C}^{d_0}\left(\mathbf{F}^d(\mathbf{b})\right)^T\mathbf{S}^{-1}\left\{\begin{pmatrix}\mathbf{F}^d(\mathbf{b}) & \mathbf{F}^{m_1}(\mathbf{b}) & \mathbf{F}^{m_2}(\mathbf{b})\end{pmatrix}\begin{pmatrix}\mathbf{d}-\mathbf{d}^0\\ \mathbf{p}^1-\mathbf{p}_0^1\\ \mathbf{p}^2-\mathbf{p}_0^2\end{pmatrix}-\mathbf{f}(\mathbf{b})\right\}\qquad(3.4.11)$$

$$\mathbf{p}^1=\mathbf{p}_0^1+\mathbf{C}_1^{p_0}\left(\mathbf{F}^{m_1}(\mathbf{b})\right)^T\mathbf{S}^{-1}\left\{\begin{pmatrix}\mathbf{F}^d(\mathbf{b}) & \mathbf{F}^{m_1}(\mathbf{b}) & \mathbf{F}^{m_2}(\mathbf{b})\end{pmatrix}\begin{pmatrix}\mathbf{d}-\mathbf{d}^0\\ \mathbf{p}^1-\mathbf{p}_0^1\\ \mathbf{p}^2-\mathbf{p}_0^2\end{pmatrix}-\mathbf{f}(\mathbf{b})\right\}\qquad(3.4.12)$$

$$\mathbf{p}^2=\mathbf{p}_0^2+\mathbf{C}_2^{p_0}\left(\mathbf{F}^{m_2}(\mathbf{b})\right)^T\mathbf{S}^{-1}\left\{\begin{pmatrix}\mathbf{F}^d(\mathbf{b}) & \mathbf{F}^{m_1}(\mathbf{b}) & \mathbf{F}^{m_2}(\mathbf{b})\end{pmatrix}\begin{pmatrix}\mathbf{d}-\mathbf{d}^0\\ \mathbf{p}^1-\mathbf{p}_0^1\\ \mathbf{p}^2-\mathbf{p}_0^2\end{pmatrix}-\mathbf{f}(\mathbf{b})\right\}\qquad(3.4.13)$$

按照一个模型函数中的同样做法，把（3.4.8)，(3.4.11）一（3.4.13）返回到连续模型中去。得

$$S_{ij}=F_{ik}^dC_{kl}^{d_0}F_{jl}^d+\frac{\partial f_i(\mathbf{d},m_1(\mathbf{x}'),m_2(\mathbf{x}'))}{\partial m_1}*C_1^{m_0}(\mathbf{x}',\mathbf{x}'')*\frac{\partial f_j(\mathbf{d},m_1(\mathbf{x}''),m_2(\mathbf{x}''))}{\partial m_1}$$
$$+\frac{\partial f_i(\mathbf{d},m_1(\mathbf{x}'),m_2(\mathbf{x}'))}{\partial m_2}*C_2^{m_0}(\mathbf{x}',\mathbf{x}'')*\frac{\partial f_j(\mathbf{d},m_1(\mathbf{x}''),m_2(\mathbf{x}''))}{\partial m_2}$$

（3.4.14）

$$F_{ij}^d(\mathbf{b})\left(d_j-d_j^0\right)+F_{ij}^{m_1}(\mathbf{b})\left(p_j^1-p_{0j}^1\right)+F_{ij}^{m_2}(\mathbf{b})\left(p_j^2-p_{0j}^2\right)=$$

$$F_{ij}^d(\mathbf{d},m_1,m_2)\left(d_j-d_j^0\right)+\frac{\partial f_i(\mathbf{d},m_1(\mathbf{x}'),m_2(\mathbf{x}'))}{\partial m_1}*\left(m_1(\mathbf{x}')-m_1^0(\mathbf{x}')\right)$$

$$+\frac{\partial f_i(\mathbf{d}, m_1(\mathbf{x}'), m_2(\mathbf{x}'))}{\partial m_2} * \left(m_2(\mathbf{x}') - m_2^0(\mathbf{x}')\right)$$

根据获得(3.2.38)同样的步骤，可得

$$d_i = d_i^0 + C_{ik}^{d_0} F_{jk}^d(\mathbf{d}, m_1, m_2) S_{jl}^{-1} \left\{ F_{ln}^d(\mathbf{d}, m_1, m_2)\left(d_n - d_n^0\right) + \frac{\partial f_l(\mathbf{d}, m_1(\mathbf{x}'), m_2(\mathbf{x}'))}{\partial m_1} * \right.$$

$$\left(m_1(\mathbf{x}') - m_1^0(\mathbf{x}')\right)$$

$$\left. + \frac{\partial f_l(\mathbf{d}, m_1(\mathbf{x}'), m_2(\mathbf{x}'))}{\partial m_2} * \left(m_2(\mathbf{x}') - m_2^0(\mathbf{x}')\right) - f_l(\mathbf{d}, m_1, m_2) \right\}$$

(3.4.15)

$$m_1(\mathbf{x}) = m_1^0(\mathbf{x}) + C_1^{m_0}(\mathbf{x}, \mathbf{x}') * \frac{\partial f_j(\mathbf{d}, m_1(\mathbf{x}'), m_2(\mathbf{x}'))}{\partial m_1} S_{jl}^{-1} \left\{ F_{ln}^d(\mathbf{d}, m_1, m_2)\left(d_n - d_n^0\right) \right.$$

$$+ \frac{\partial f_l(\mathbf{d}, m_1(\mathbf{x}'), m_2(\mathbf{x}'))}{\partial m_1} * \left(m_1(\mathbf{x}') - m_1^0(\mathbf{x}')\right) \tag{3.4.16}$$

$$\left. + \frac{\partial f_l(\mathbf{d}, m_1(\mathbf{x}'), m_2(\mathbf{x}'))}{\partial m_2} * \left(m_2(\mathbf{x}') - m_2^0(\mathbf{x}')\right) - f_l(\mathbf{d}, m_1, m_2) \right\}$$

$$m_2(\mathbf{x}) = m_2^0(\mathbf{x}) + C_2^{m_0}(\mathbf{x}, \mathbf{x}') * \frac{\partial f_j(\mathbf{d}, m_1(\mathbf{x}'), m_2(\mathbf{x}'))}{\partial m_2} S_{jl}^{-1} \left\{ F_{ln}^d(\mathbf{d}, m_1, m_2)\left(d_n - d_n^0\right) \right.$$

$$+ \frac{\partial f_l(\mathbf{d}, m_1(\mathbf{x}'), m_2(\mathbf{x}'))}{\partial m_1} * \left(m_1(\mathbf{x}') - m_1^0(\mathbf{x}')\right) \tag{3.4.17}$$

$$\left. + \frac{\partial f_l(\mathbf{d}, m_1(\mathbf{x}'), m_2(\mathbf{x}'))}{\partial m_2} * \left(m_2(\mathbf{x}') - m_2^0(\mathbf{x}')\right) - f_l(\mathbf{d}, m_1, m_2) \right\}$$

由此可以推广到一共有 k 个模型函数的情况。假设先验模型函数之间相互独立，第 α 个模型函数的先验协方差函数为 $C_\alpha^{m_0}(\mathbf{x}, \mathbf{x}')$。

1) 一般情况

正问题： $f_i(\mathbf{d}, m_1(\mathbf{x}), \cdots, m_k(\mathbf{x})) = 0$，$(i = 1, \cdots, n)$ (3.4.18)

反问题的解：

设第 α 个模型函数 $m_\alpha(\mathbf{x})$ 的先验均值解为 $m_\alpha^0(\mathbf{x})$，关于 $m_\alpha^0(\mathbf{x})$ 的先验协方差函数为 $C_\alpha^{m_0}(\mathbf{x}, \mathbf{x}')$。则

$$m_\alpha(\mathbf{x}) = m_\alpha^0(\mathbf{x}) + C_\alpha^{m_0}(\mathbf{x},\mathbf{x}') * \frac{\partial f_j(\mathbf{d}, m_1(\mathbf{x}'), m_2(\mathbf{x}'), \cdots, m_k(\mathbf{x}'))}{\partial m_\alpha} S_{jl}^{-1}$$

$$\left\{F_{ln}^d(\mathbf{d}, m_1, m_2, \cdots, m_k)\left(d_n - d_n^0\right)\right. \tag{3.4.19}$$

$$\left. + \sum_{\beta=1}^{k} \frac{\partial f_l(\mathbf{d}, m_1(\mathbf{x}'), m_2(\mathbf{x}'), \cdots, m_k(\mathbf{x}'))}{\partial m_\beta} * \left(m_\beta(\mathbf{x}') - m_\beta^0(\mathbf{x}')\right) - f_l(\mathbf{d}, m_1, m_2 \cdots, m_k)\right\}$$

$$d_\alpha = d_\alpha^0 + C_{\alpha k}^{d_0} F_{jk}^d(\mathbf{d}, m_1, m_2, \cdots, m_k) S_{jl}^{-1}\left\{F_{ln}^d(\mathbf{d}, m_1, m_2, \cdots, m_k)\left(d_n - d_n^0\right)\right.$$

$$\left. + \sum_{\beta=1}^{k} \frac{\partial f_l(\mathbf{d}, m_1(\mathbf{x}'), m_2(\mathbf{x}'), \cdots, m_k(\mathbf{x}'))}{\partial m_\beta} * \left(m_\beta(\mathbf{x}') - m_\beta^0(\mathbf{x}')\right) - f_l(\mathbf{d}, m_1, m_2 \cdots, m_k)\right\} \tag{3.4.20}$$

$$S_{ij} = F_{in}^d C_{nl}^{d_0} F_{jl}^d$$

$$+ \sum_{\alpha=1}^{k} \frac{\partial f_i(\mathbf{d}, m_1(\mathbf{x}'), \cdots, m_k(\mathbf{x}'))}{\partial m_\alpha} * C_\alpha^{m_0}(\mathbf{x}', \mathbf{x}'') * \frac{\partial f_j(\mathbf{d}, m_1(\mathbf{x}''), \cdots, m_k(\mathbf{x}'))}{\partial m_\alpha} \tag{3.4.21}$$

对于非线性系统，（3.4.19）(3.4.20)是迭代解。

线性系统的解

正问题：$f_i(\mathbf{d}, m_1, \cdots, m_k) = F_{ij}^d d_j + \sum_{\alpha=1}^{k} g_i^\alpha(\mathbf{x}) * m_\alpha(\mathbf{x}) = 0$，$(i = 1, \cdots, n)$ （3.4.22）

其中算子 $g_i^\alpha = \dfrac{\partial f_i}{\partial m_\alpha}$ 不含有模型函数和数据分量。

反问题的解：

$$\tilde{m}_\alpha(\mathbf{x}) = m_\alpha^0(\mathbf{x}) - C_\alpha^{m_0}(\mathbf{x},\mathbf{x}') * g_j^\alpha(\mathbf{x}') S_{jl}^{-1}\left\{F_{ln}^d d_n^0 + \sum_{\beta=1}^{k} g_l^\beta(\mathbf{x}') * m_\beta^0(\mathbf{x}')\right\} \tag{3.4.23}$$

$$\tilde{d}_\alpha = d_\alpha^0 - C_{\alpha k}^{d_0} F_{jk}^d S_{jl}^{-1}\left\{F_{ln}^d d_n^0 + \sum_{\beta=1}^{k} g_l^\beta(\mathbf{x}') * m_\beta^0(\mathbf{x}')\right\} \tag{3.4.24}$$

$$S_{ij} = F_{in}^d C_{nl}^{d_0} F_{jl}^d + \sum_{\alpha=1}^{k} g_i^\alpha(\mathbf{x}') * C_\alpha^{m_0}(\mathbf{x}', \mathbf{x}'') * g_j^\alpha(\mathbf{x}'') \tag{3.4.25}$$

2) 数据可以用模型函数显式表示

正问题：$f(\mathbf{d}, m_1(\mathbf{x}), \cdots, m_k(\mathbf{x})) = d_i - g_i(m_1(\mathbf{x}), \cdots, m_k(\mathbf{x})) = 0$，$(i = 1, \cdots, n)$ (3.4.26)

反问题的解：

$$\frac{\partial f_i}{\partial m_j} = -\frac{\partial g_i}{\partial m_j}$$

$$S_{ij}=C_{ij}^{d_0}+\sum_{\alpha=1}^{k}\frac{\partial g_i(m_1(\mathbf{x}'),\cdots,m_k(\mathbf{x}'))}{\partial m_\alpha}*C_\alpha^{m_0}(\mathbf{x}',\mathbf{x}'')*\frac{\partial g_j(m_1(\mathbf{x}''),\cdots,m_k(\mathbf{x}''))}{\partial m_\alpha}$$

(3.4.27)

$$d_i=d_i^0-C_{ij}^{d_0}S_{jl}^{-1}\left\{(d_l^0-d_l)+\sum_{\alpha=1}^{k}\frac{\partial g_l(m_1(\mathbf{x}'),\cdots,\ m_k(\mathbf{x}'))}{\partial m_\alpha}*(m_\alpha(\mathbf{x}')-m_\alpha^0(\mathbf{x}'))\right\}$$

(3.3.4.28)

$$m_\alpha(\mathbf{x})=m_\alpha^0(\mathbf{x})+\left(C_\alpha^{m_0}(\mathbf{x},\mathbf{x}')*\frac{\partial g_j(m_1(\mathbf{x}'),\cdots,m_k(\mathbf{x}'))}{\partial m_\alpha}\right)$$
$$S_{jl}^{-1}\left\{(d_l^0-d_l)+\sum_{\alpha=1}^{k}\frac{\partial g_l(m_1(\mathbf{x}'),\cdots,m_k(\mathbf{x}'))}{\partial m_\alpha}*(m_\alpha(\mathbf{x}')-m_\alpha^0(\mathbf{x}'))\right\}$$
(3.4.29)

对于非线性系统，（3.4.28）(3.4.29)是迭代解.

线性系统的解

正问题：$f_i(\mathbf{d},m_1,\cdots,m_k)=d_i-\sum_{\alpha=1}^{k}g_i^\alpha(\mathbf{x})*m_\alpha(\mathbf{x})=0$ (3.4.30)

反问题解：

$$S_{ij}=C_{ij}^{d_0}+\sum_{\alpha=1}^{k}g_i^\alpha(\mathbf{x}')*C_\alpha^{m_0}(\mathbf{x}',\mathbf{x}'')*g_j^\alpha(\mathbf{x}'')\qquad(3.4.31)$$

$$\tilde{d}_i=d_i^0-C_{ij}^{d_0}S_{jl}^{-1}\left\{d_l^0-\sum_{\alpha=1}^{k}g_l^\alpha(\mathbf{x}')*m_\alpha^0(\mathbf{x}')\right\}=\sum_{\alpha=1}^{k}g_i^\alpha(\mathbf{x}')*\tilde{m}_\alpha(\mathbf{x}')\qquad(3.4.32)$$

$$\tilde{m}_\alpha(\mathbf{x})=m_\alpha^0(\mathbf{x})+\left(g_i^\alpha(\mathbf{x}')*C_\alpha^{m_0}(\mathbf{x},\mathbf{x}')\right)S_{ij}^{-1}\left\{d_j^0-\sum_{\beta=1}^{k}g_j^\beta(\mathbf{x}')*m_\beta^0(\mathbf{x}')\right\}\qquad(3.4.33)$$

3) 无数据

此情况，可以用来求微分方程的解。

正问题： $f_i(\ m_1(\mathbf{x}),\cdots,m_k(\mathbf{x}))=0$，$(i=1,\cdots,n)$

反问题解：

$$m_\alpha(\mathbf{x})=m_\alpha^0(\mathbf{x})+C_\alpha^{m_0}(\mathbf{x},\mathbf{x}')*\frac{\partial f_j(m_1(\mathbf{x}'),m_2(\mathbf{x}'),\cdots,m_k(\mathbf{x}'))}{\partial m_\alpha}S_{jl}^{-1}$$
$$\left\{\sum_{\beta=1}^{k}\frac{\partial f_l(m_1(\mathbf{x}'),m_2(\mathbf{x}'),\cdots,m_k(\mathbf{x}'))}{\partial m_\beta}*\left(m_\beta(\mathbf{x}')-m_\beta^0(\mathbf{x}')\right)-f_l(m_1,m_2\cdots,m_k)\right\}$$
(3.4.34)

$$S_{ij}=\sum_{\alpha=1}^{k}\frac{\partial f_i(m_1(\mathbf{x}'),\cdots,m_k(\mathbf{x}'))}{\partial m_\alpha}*C_\alpha^{m_0}(\mathbf{x}',\mathbf{x}'')*\frac{\partial f_j(m_1(\mathbf{x}''),\cdots,m_k(\mathbf{x}'))}{\partial m_\alpha}\qquad(3.4.35)$$

非线性系统必须迭代。

4) 多模型函数反演的分辨与误差

先讨论两个模型函数的情况。设数据可以用模型函数显式表示。

正问题为：$d_i - g_i(m_1(\mathbf{x}), m_2(\mathbf{x})) = 0$

模型函数及相关的量按照(3.4.2) －（3.4.10）表示后，设矩阵$\mathbf{G}_1$的元素由$\dfrac{\partial g_i}{\partial p_{1j}}$构成，矩阵$\mathbf{G}_2$的元素由$\dfrac{\partial g_i}{\partial p_{2j}}$构成，则$\mathbf{G} = \begin{pmatrix}\mathbf{G}_1 & \mathbf{G}_2\end{pmatrix}$。由于先验模型函数相互独立，则

$\mathbf{C}_0^{mm} = \begin{pmatrix}\mathbf{C}_1^{p_0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_2^{p_0}\end{pmatrix}$，令后验协方差矩阵为$\mathbf{C}^{mm} = \begin{pmatrix}\mathbf{C}^{p_1p_1} & \mathbf{C}^{p_1p_2} \\ \mathbf{C}^{p_2p_1} & \mathbf{C}^{p_2p_2}\end{pmatrix}$，根据（3.1.58）后验协方差矩阵可表示为

$$
\begin{aligned}
\begin{pmatrix}\mathbf{C}^{p_1p_1} & \mathbf{C}^{p_1p_2} \\ \mathbf{C}^{p_2p_1} & \mathbf{C}^{p_2p_2}\end{pmatrix} &= \begin{pmatrix}\mathbf{C}_1^{p_0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_2^{p_0}\end{pmatrix} - \begin{pmatrix}\mathbf{C}_1^{p_0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_2^{p_0}\end{pmatrix}\begin{pmatrix}\mathbf{G}_1^T \\ \mathbf{G}_2^T\end{pmatrix}\mathbf{S}^{-1}\begin{pmatrix}\mathbf{G}_1 & \mathbf{G}_2\end{pmatrix}\begin{pmatrix}\mathbf{C}_1^{p_0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_2^{p_0}\end{pmatrix} \\
&= \begin{pmatrix}\mathbf{C}_1^{p_0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_2^{p_0}\end{pmatrix} - \begin{pmatrix}\mathbf{C}_1^{p_0}\mathbf{G}_1^T \\ \mathbf{C}_2^{p_0}\mathbf{G}_2^T\end{pmatrix}\mathbf{S}^{-1}\begin{pmatrix}\mathbf{G}_1\mathbf{C}_1^{p_0} & \mathbf{G}_2\mathbf{C}_2^{p_0}\end{pmatrix} \\
&= \begin{pmatrix}\mathbf{C}_1^{p_0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_2^{p_0}\end{pmatrix} - \begin{pmatrix}\mathbf{C}_1^{p_0}\mathbf{G}_1^T\mathbf{S}^{-1}\mathbf{G}_1\mathbf{C}_1^{p_0} & \mathbf{C}_1^{p_0}\mathbf{G}_1^T\mathbf{S}^{-1}\mathbf{G}_2\mathbf{C}_2^{p_0} \\ \mathbf{C}_2^{p_0}\mathbf{G}_2^T\mathbf{S}^{-1}\mathbf{G}_1\mathbf{C}_1^{p_0} & \mathbf{C}_2^{p_0}\mathbf{G}_2^T\mathbf{S}^{-1}\mathbf{G}_2\mathbf{C}_2^{p_0}\end{pmatrix} \\
&= \begin{pmatrix}\mathbf{C}_1^{p_0} - \mathbf{C}_1^{p_0}\mathbf{G}_1^T\mathbf{S}^{-1}\mathbf{G}_1\mathbf{C}_1^{p_0} & -\mathbf{C}_1^{p_0}\mathbf{G}_1^T\mathbf{S}^{-1}\mathbf{G}_2\mathbf{C}_2^{p_0} \\ -\mathbf{C}_2^{p_0}\mathbf{G}_2^T\mathbf{S}^{-1}\mathbf{G}_1\mathbf{C}_1^{p_0} & \mathbf{C}_2^{p_0} - \mathbf{C}_2^{p_0}\mathbf{G}_2^T\mathbf{S}^{-1}\mathbf{G}_2\mathbf{C}_2^{p_0}\end{pmatrix}
\end{aligned}
$$

即

$$
\begin{aligned}
\mathbf{C}^{p_1p_1} &= \mathbf{C}_1^{p_0} - \mathbf{C}_1^{p_0}\mathbf{G}_1^T\mathbf{S}^{-1}\mathbf{G}_1\mathbf{C}_1^{p_0} \\
\mathbf{C}^{p_2p_2} &= \mathbf{C}_2^{p_0} - \mathbf{C}_2^{p_0}\mathbf{G}_2^T\mathbf{S}^{-1}\mathbf{G}_2\mathbf{C}_2^{p_0} \\
\mathbf{C}^{p_1p_2} &= -\mathbf{C}_1^{p_0}\mathbf{G}_1^T\mathbf{S}^{-1}\mathbf{G}_2\mathbf{C}_2^{p_0} \\
\mathbf{C}^{p_2p_1} &= -\mathbf{C}_2^{p_0}\mathbf{G}_2^T\mathbf{S}^{-1}\mathbf{G}_1\mathbf{C}_1^{p_0}
\end{aligned}
$$

返回到连续模型，根据求(3.2.57)相同的步骤，就得到两个模型函数时后验协方差及交叉协方差函数

$$
S_{ij} = C_{ij}^{d_0} + \sum_{\alpha=1}^{2}\frac{\partial g_i(m_1(\mathbf{x}'), m_2(\mathbf{x}'))}{\partial m_\alpha} * C_\alpha^{m_0}(\mathbf{x}', \mathbf{x}'') * \frac{\partial g_j(m_1(\mathbf{x}''), m_2(\mathbf{x}''))}{\partial m_\alpha}
$$

第一个模型函数$m_1(\mathbf{x})$的后验协方差函数为

$$C_{11}^{m}(\mathbf{x},\mathbf{x}')=C_{1}^{m_0}(\mathbf{x},\mathbf{x}')$$
$$-C_{1}^{m_0}(\mathbf{x},\mathbf{x}'')*\frac{\partial g_i(m_1(\mathbf{x}''),m_2(\mathbf{x}''))}{\partial m_1}S_{ij}^{-1}\frac{\partial g_j(m_1(\mathbf{x}'''),m_2(\mathbf{x}'''))}{\partial m_1}*C_{1}^{m_0}(\mathbf{x}''',\mathbf{x}')$$

第二个模型函数 $m_2(\mathbf{x})$ 的后验协方差函数为

$$C_{22}^{m}(\mathbf{x},\mathbf{x}')=C_{2}^{m_0}(\mathbf{x},\mathbf{x}')$$
$$-C_{2}^{m_0}(\mathbf{x},\mathbf{x}'')*\frac{\partial g_i(m_1(\mathbf{x}''),m_2(\mathbf{x}''))}{\partial m_2}S_{ij}^{-1}\frac{\partial g_j(m_1(\mathbf{x}'''),m_2(\mathbf{x}'''))}{\partial m_2}*C_{2}^{m_0}(\mathbf{x}''',\mathbf{x}')$$

两个模型函数的交叉协方差为

$$C_{12}^{m}(\mathbf{x},\mathbf{x}')=-C_{1}^{m_0}(\mathbf{x},\mathbf{x}'')*\frac{\partial g_i(m_1(\mathbf{x}''),m_2(\mathbf{x}''))}{\partial m_1}S_{ij}^{-1}\frac{\partial g_j(m_1(\mathbf{x}'''),m_2(\mathbf{x}'''))}{\partial m_2}*C_{2}^{m_0}(\mathbf{x}''',\mathbf{x}')$$

$$C_{21}^{m}(\mathbf{x},\mathbf{x}')=-C_{2}^{m_0}(\mathbf{x},\mathbf{x}'')*\frac{\partial g_i(m_1(\mathbf{x}''),m_2(\mathbf{x}''))}{\partial m_2}S_{ij}^{-1}\frac{\partial g_j(m_1(\mathbf{x}'''),m_2(\mathbf{x}'''))}{\partial m_1}*C_{1}^{m_0}(\mathbf{x}''',\mathbf{x}')$$

由 $\mathbf{S}$ 的对称性推知

$$C_{12}^{m}(\mathbf{x},\mathbf{x}')=C_{21}^{m}(\mathbf{x},\mathbf{x}')$$

根据以上的分析很容易证明，当模型函数一共有 k 个时，第 α 个模型函数的后验协方差函数为

$$C_{\alpha\alpha}^{m}(\mathbf{x},\mathbf{x}')=C_{\alpha}^{m_0}(\mathbf{x},\mathbf{x}')$$
$$-C_{\alpha}^{m_0}(\mathbf{x},\mathbf{x}'')*\frac{\partial g_i(m_1(\mathbf{x}''),\cdots,m_k(\mathbf{x}''))}{\partial m_\alpha}S_{ij}^{-1}\frac{\partial g_j(m_1(\mathbf{x}'''),\cdots,m_k(\mathbf{x}'''))}{\partial m_\alpha}*C_{\alpha}^{m_0}(\mathbf{x}''',\mathbf{x}')$$

（3.4.36）

第 α 个模型函数与第 β 个模型函数的后验交叉协方差函数为

$$C_{\alpha\beta}^{m}(\mathbf{x},\mathbf{x}')=-C_{\alpha}^{m_0}(\mathbf{x},\mathbf{x}'')*\frac{\partial g_i(m_1(\mathbf{x}''),\cdots,m_k(\mathbf{x}''))}{\partial m_\alpha}$$
$$S_{ij}^{-1}\frac{\partial g_j(m_1(\mathbf{x}'''),\cdots,m_k(\mathbf{x}'''))}{\partial m_\beta}*C_{\beta}^{m_0}(\mathbf{x}''',\mathbf{x}') \quad (3.4.37)$$

$$S_{ij}=C_{ij}^{d_0}+\sum_{\alpha=1}^{k}\frac{\partial g_i(m_1(\mathbf{x}'),\cdots,m_k(\mathbf{x}'))}{\partial m_\alpha}*C_{\alpha}^{m_0}(\mathbf{x}',\mathbf{x}'')*\frac{\partial g_j(m_1(\mathbf{x}''),\cdots,m_k(\mathbf{x}''))}{\partial m_\alpha} \quad (3.4.38)$$

对于非线性问题，（3.4.36）－（3.4.38）中的模型函数取最后一次的迭代值。当 $\mathbf{x}=\mathbf{x}'$ 时，$C_{\alpha\alpha}^{m}(\mathbf{x},\mathbf{x})$ 表示第 α 个模型函数在 $\mathbf{x}$ 点的方差。当 $\mathbf{x}\neq\mathbf{x}'$ 时，取定 $\mathbf{x}$，$C_{\alpha\alpha}^{m}(\mathbf{x},\mathbf{x}')$ 表示模型

函数在 $\mathbf{x}$ 点的分辨。由（3.4.37）可知，虽然假定了先验模型函数之间相互独立，但后验模型函数之间是有可能相关的。

3.5 混合问题

描述一个物理系统有时仅用离散参数还不够，同时还需要用连续函数。这是模型参数的混合问题。比如，利用走时资料进行震源参数反演，这是离散模型参数的反演问题，有4个离散参数，震源的坐标和发震时刻。如果把地震波速度和震源参数联合反演，速度模型用连续函数表示，那么反演就成了混合问题。又比如，面波波形反演，介质速度模型的描述在有些层可用离散参数，在有些深度上可以用连续模型（比如上地幔过渡带），这也是模型参数的混合问题。

为简单起见，首先考虑一组离散模型参数和一个模型函数的情况。数据是离散的，d_i，$(i=1,\cdots,n)$， 先验均值为 d_i^0，先验协方差矩阵为 $C_{ij}^{d_0}$。离散模型参数用 $\mathbf{h}=(h_1 \quad \cdots \quad h_s)^T$ 表示，具有先验均值 $\mathbf{h}^0=(h_1^0 \quad \cdots \quad h_s^0)^T$ 和先验协方差矩阵 $\mathbf{C}^{h_0}$。模型函数用 $m(\mathbf{x})$ 表示，具有先验均值 $m^0(\mathbf{x})$ 和关于 $m^0(\mathbf{x})$ 的先验协方差函数 $C^{m_0}(\mathbf{x},\mathbf{x}')$。设先验均值 d_i^0 与先验模型参数及函数之间不相关，离散模型参数与模型函数之间也不相关。正问题为

$$f_i(\mathbf{d},\mathbf{h},m(\mathbf{x}))=0\text{，}\quad (i=1,\cdots,n) \tag{3.5.1}$$

如果模型函数像（3.2.14）那样用基函数展开，其坐标为 p_i，则正问题可表述为离散模型的形式

$$f_i(d_1,\cdots,d_n,h_1,\cdots,h_s,p_1,p_2,\cdots)=0\text{，}\ (i=1,\cdots,n) \tag{3.5.2}$$

先验均值函数 $m^0(\mathbf{x})$ 用（3.2.16）展开，其坐标为 p_i^0，先验协方差函数 $C^{m_0}(\mathbf{x},\mathbf{x}')$ 像（3.2.18）那样用矩阵表示，$C_{ij}^{p_0}$ 。令

$$\mathbf{b}=(d_1\cdots d_n \ h_1\cdots h_s \quad p_1 \quad p_2\cdots)^T \tag{3.5.3}$$

$$\mathbf{b}_0=(d_1^0 \ d_2^0 \cdots d_n^0 \ h_1^0 \ h_2^0\cdots p_1^0 \ p_2^0\cdots)^T \tag{3.5.4}$$

则数据和模型分量的先验协方差矩阵分别为 $C_{ij}^{d_0}$，$C_{ij}^{h_0}$，$C_{ij}^{p_0}$，总的先验协方差矩阵为

$$\mathbf{C}_0=\begin{pmatrix}\mathbf{C}^{d_0} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{C}^{h_0} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{C}^{p_0}\end{pmatrix} \tag{3.5.5}$$

令

$$\mathbf{S}(\mathbf{b})=\mathbf{F}(\mathbf{b})\mathbf{C}_0\mathbf{F}^T(\mathbf{b})$$

$$=\begin{pmatrix}\mathbf{F}^d(\mathbf{b}) & \mathbf{F}^h(\mathbf{b}) & \mathbf{F}^p(\mathbf{b})\end{pmatrix}\begin{pmatrix}\mathbf{C}^{d_0} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{C}^{h_0} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{C}^{p_0}\end{pmatrix}\begin{pmatrix}\left(\mathbf{F}^d(\mathbf{b})\right)^T\\ \left(\mathbf{F}^h(\mathbf{b})\right)^T\\ \left(\mathbf{F}^p(\mathbf{b})\right)^T\end{pmatrix}$$

$$=\mathbf{F}^d(\mathbf{b})\mathbf{C}^{d_0}\left(\mathbf{F}^d(\mathbf{b})\right)^T+\mathbf{F}^h(\mathbf{b})\mathbf{C}^{h_0}\left(\mathbf{F}^h(\mathbf{b})\right)^T+\mathbf{F}^p(\mathbf{b})\mathbf{C}^{p_0}\left(\mathbf{F}^p(\mathbf{b})\right)^T \qquad (3.5.6)$$

$$F_{ij}(\mathbf{b})=\frac{\partial f_i(\mathbf{b})}{\partial b_j}，\quad F_{ij}^d(\mathbf{b})=\frac{\partial f_i(\mathbf{b})}{\partial d_j}*I$$

$$F_{ij}^h(\mathbf{b})=\frac{\partial f_i(\mathbf{b})}{\partial h_j}*I，\quad F_{ij}^p(\mathbf{b})=\frac{\partial f_i(\mathbf{b})}{\partial p_j}=\frac{\partial f_i(\mathbf{d},\mathbf{h},m(\mathbf{x}))}{\partial m}*\varphi_j(\mathbf{x}) \qquad (3.5.7)$$

根据（3.1.20）得

$$\mathbf{b}=\mathbf{b}_0+\mathbf{C}_0\mathbf{F}^T(\mathbf{b})\left(\mathbf{F}(\mathbf{b})\mathbf{C}_0\mathbf{F}^T(\mathbf{b})\right)^{-1}\left\{\mathbf{F}(\mathbf{b})(\mathbf{b}-\mathbf{b}_0)-\mathbf{f}(\mathbf{b})\right\}$$

或

$$\begin{pmatrix}\mathbf{d}\\ \mathbf{h}\\ \mathbf{p}\end{pmatrix}=\begin{pmatrix}\mathbf{d}^0\\ \mathbf{h}^0\\ \mathbf{p}^0\end{pmatrix}+\begin{pmatrix}\mathbf{C}^{d_0} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{C}^{h_0} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{C}^{p_0}\end{pmatrix}\begin{pmatrix}\left(\mathbf{F}^d(\mathbf{b})\right)^T\\ \left(\mathbf{F}^h(\mathbf{b})\right)^T\\ \left(\mathbf{F}^p(\mathbf{b})\right)^T\end{pmatrix}$$

$$\mathbf{S}^{-1}\left\{\begin{pmatrix}\mathbf{F}^d(\mathbf{b}) & \mathbf{F}^h(\mathbf{b}) & \mathbf{F}^p(\mathbf{b})\end{pmatrix}\begin{pmatrix}\mathbf{d}-\mathbf{d}^0\\ \mathbf{h}-\mathbf{h}^0\\ \mathbf{p}-\mathbf{p}^0\end{pmatrix}-\mathbf{f}(\mathbf{b})\right\}$$

所以

$$\mathbf{d}=\mathbf{d}^0+\mathbf{C}^{d_0}\left(\mathbf{F}^d(\mathbf{b})\right)^T\mathbf{S}^{-1}\left\{\begin{pmatrix}\mathbf{F}^d(\mathbf{b}) & \mathbf{F}^h(\mathbf{b}) & \mathbf{F}^p(\mathbf{b})\end{pmatrix}\begin{pmatrix}\mathbf{d}-\mathbf{d}^0\\ \mathbf{h}-\mathbf{h}^{\mathbf{0}}\\ \mathbf{p}-\mathbf{p}^0\end{pmatrix}-\mathbf{f}(\mathbf{b})\right\} \qquad (3.5.8)$$

$$\mathbf{h}=\mathbf{h}^0+\mathbf{C}^{h_0}\left(\mathbf{F}^h(\mathbf{b})\right)^T\mathbf{S}^{-1}\left\{\begin{pmatrix}\mathbf{F}^d(\mathbf{b}) & \mathbf{F}^h(\mathbf{b}) & \mathbf{F}^p(\mathbf{b})\end{pmatrix}\begin{pmatrix}\mathbf{d}-\mathbf{d}^0\\ \mathbf{h}-\mathbf{h}^{\mathbf{0}}\\ \mathbf{p}-\mathbf{p}^0\end{pmatrix}-\mathbf{f}(\mathbf{b})\right\} \qquad (3.5.9)$$

$$\mathbf{p}=\mathbf{p}^0+\mathbf{C}^{p_0}\left(\mathbf{F}^p(\mathbf{b})\right)^T\mathbf{S}^{-1}\left\{\begin{pmatrix}\mathbf{F}^d(\mathbf{b}) & \mathbf{F}^h(\mathbf{b}) & \mathbf{F}^p(\mathbf{b})\end{pmatrix}\begin{pmatrix}\mathbf{d}-\mathbf{d}^0\\ \mathbf{h}-\mathbf{h}^{\mathbf{0}}\\ \mathbf{p}-\mathbf{p}^0\end{pmatrix}-\mathbf{f}(\mathbf{b})\right\} \qquad (3.5.10)$$

把（3.5.6),（3.5.8）一（3.5.10）返回到连续模型中去。得

$$S_{ij}=F_{ik}^{d}C_{kl}^{d_0}F_{jl}^{d}+F_{ik}^{h}C_{kl}^{h_0}F_{jl}^{h}+\frac{\partial f_i(\mathbf{d},m(\mathbf{x}'))}{\partial m}*C^{m_0}(\mathbf{x}',\mathbf{x}'')*\frac{\partial f_j(\mathbf{d},m(\mathbf{x}''))}{\partial m} \tag{3.5.11}$$

$$\begin{aligned}&F_{ij}^{d}(\mathbf{b})\left(d_j-d_j^0\right)+F_{ij}^{h}(\mathbf{b})\left(h_j-h_j^0\right)+F_{ij}^{p}(\mathbf{b})\left(p_j-p_j^0\right)=\\&F_{ij}^{d}(\mathbf{d},\mathbf{h},m)\left(d_j-d_j^0\right)+F_{ij}^{h}(\mathbf{d},\mathbf{h},m)\left(h_j-h_j^0\right)+\frac{\partial f_i(\mathbf{d},\mathbf{h},m(\mathbf{x}'))}{\partial m}*\left(m(\mathbf{x}')-m^0(\mathbf{x}')\right)\end{aligned} \tag{3.5.12}$$

$$\begin{aligned}d_i=d_i^0+C_{ik}^{d_0}F_{jk}^{d}(\mathbf{d},\mathbf{h},m)S_{jl}^{-1}&\Big\{F_{ln}^{d}(\mathbf{d},\mathbf{h},m)\left(d_n-d_n^0\right)+F_{ln}^{h}(\mathbf{d},\mathbf{h},m)\left(h_n-h_n^0\right)\\&+\frac{\partial f_l(\mathbf{d},\mathbf{h},m(\mathbf{x}'))}{\partial m}*\left(m(\mathbf{x}')-m^0(\mathbf{x}')\right)-f_l(\mathbf{d},m_1,m_2)\Big\}\end{aligned} \tag{3.5.13}$$

$$\begin{aligned}h_i=h_i^0+C_{ik}^{h_0}F_{jk}^{h}(\mathbf{d},\mathbf{h},m)S_{jl}^{-1}&\Big\{F_{ln}^{d}(\mathbf{d},\mathbf{h},m)\left(d_n-d_n^0\right)+F_{ln}^{h}(\mathbf{d},\mathbf{h},m)\left(h_n-h_n^0\right)\\&+\frac{\partial f_l(\mathbf{d},\mathbf{h},m(\mathbf{x}'))}{\partial m}*\left(m(\mathbf{x}')-m^0(\mathbf{x}')\right)-f_l(\mathbf{d},m_1,m_2)\Big\}\end{aligned} \tag{3.5.14}$$

$$\begin{aligned}m(\mathbf{x})=m^0(\mathbf{x})+C^{m_0}(\mathbf{x},\mathbf{x}')*\frac{\partial f_j(\mathbf{d},\mathbf{h},m(\mathbf{x}'))}{\partial m}S_{jl}^{-1}&\Big\{F_{ln}^{d}(\mathbf{d},\mathbf{h},m)\left(d_n-d_n^0\right)\\+F_{ln}^{h}(\mathbf{d},\mathbf{h},m)\left(h_n-h_n^0\right)+\frac{\partial f_l(\mathbf{d},\mathbf{h},m(\mathbf{x}'))}{\partial m}*\left(m(\mathbf{x}')-m^0(\mathbf{x}')\right)-f_l(\mathbf{d},m_1,m_2)\Big\}\end{aligned} \tag{3.5.15}$$

由此可以推广到一共有 k 个模型函数的情况。假设先验模型函数之间相互独立，第 α 个模型函数的先验协方差函数为 $C_{\alpha}^{m_0}(\mathbf{x},\mathbf{x}')$。

1) 一般情况

正问题： $f_i(\mathbf{d},\mathbf{p},m_1(\mathbf{x}),\cdots,m_k(\mathbf{x}))=0$，$(i=1,\cdots,n)$ (3.5.16)

反问题的解：

设第 α 个模型函数 $m_\alpha(\mathbf{x})$ 的先验均值解为 $m_\alpha^0(\mathbf{x})$，关于 $m_\alpha^0(\mathbf{x})$ 的先验协方差函数为 $C_{\alpha}^{m_0}(\mathbf{x},\mathbf{x}')$。则

$$\begin{aligned}h_i=h_i^0+C_{ik}^{h_0}F_{jk}^{h}(\mathbf{d},\mathbf{h},m)S_{jl}^{-1}&\Big\{F_{ln}^{d}(\mathbf{d},\mathbf{h},m)\left(d_n-d_n^0\right)+F_{ln}^{h}(\mathbf{d},\mathbf{h},m)\left(h_n-h_n^0\right)\\+\sum_{\beta=1}^{k}\frac{\partial f_l(\mathbf{d},\mathbf{h},m_1(\mathbf{x}'),\cdots,m_k(\mathbf{x}'))}{\partial m_\beta}&*\left(m_\beta(\mathbf{x}')-m_\beta^0(\mathbf{x}')\right)-f_l(\mathbf{d},\mathbf{h},m_1,\cdots,m_k)\Big\}\end{aligned} \tag{3.5.17}$$

$$m_\alpha(\mathbf{x}) = m_\alpha^0(\mathbf{x}) + C_\alpha^{m_0}(\mathbf{x},\mathbf{x}') * \frac{\partial f_j(\mathbf{d},\mathbf{h},m_1(\mathbf{x}'),m_2(\mathbf{x}'),\cdots,m_k(\mathbf{x}'))}{\partial m_\alpha} S_{jl}^{-1}$$
$$\Big\{F_{ln}^d(\mathbf{d},\mathbf{h},m_1,\cdots,m_k)\left(d_n - d_n^0\right) + F_{ln}^h(\mathbf{d},\mathbf{h},m_1,\cdots,m_k)\left(h_n - h_n^0\right) \quad (3.5.18)$$
$$+\sum_{\beta=1}^{k}\frac{\partial f_l(\mathbf{d},\mathbf{h},m_1(\mathbf{x}'),\cdots,m_k(\mathbf{x}'))}{\partial m_\beta} * \left(m_\beta(\mathbf{x}') - m_\beta^0(\mathbf{x}')\right) - f_l(\mathbf{d},\mathbf{h},m_1,\cdots,m_k)\Big\}$$

$$d_\alpha = d_\alpha^0 + C_{\alpha k}^{d_0} F_{jk}^d(\mathbf{d},\mathbf{h},m_1,m_2,\cdots,m_k) S_{jl}^{-1}\Big\{F_{ln}^d(\mathbf{d},\mathbf{h},m_1,m_2,\cdots,m_k)\left(d_n - d_n^0\right)$$
$$+ F_{ln}^h(\mathbf{d},\mathbf{h},m_1,m_2,\cdots,m_k)\left(h_n - h_n^0\right) \quad (3.5.19)$$
$$+\sum_{\beta=1}^{k}\frac{\partial f_l(\mathbf{d},\mathbf{h},m_1(\mathbf{x}'),\cdots,m_k(\mathbf{x}'))}{\partial m_\beta} * \left(m_\beta(\mathbf{x}') - m_\beta^0(\mathbf{x}')\right) - f_l(\mathbf{d},\mathbf{h},m_1,\cdots,m_k)\Big\}$$

$$S_{ij} = F_{in}^d C_{nl}^{d_0} F_{jl}^d + F_{in}^h C_{nl}^{h_0} F_{jl}^h$$
$$+\sum_{\alpha=1}^{k}\frac{\partial f_i(\mathbf{d},\mathbf{h},m_1(\mathbf{x}'),\cdots,m_k(\mathbf{x}'))}{\partial m_\alpha} * C_\alpha^{m_0}(\mathbf{x}',\mathbf{x}'') * \frac{\partial f_j(\mathbf{d},\mathbf{h},m_1(\mathbf{x}''),\cdots,m_k(\mathbf{x}''))}{\partial m_\alpha} \quad (3.5.20)$$

对于非线性系统，（3.5.17）(3.5.20)是迭代解。

线性系统的解

正问题：$f_i(\mathbf{d},\mathbf{h},m_1,\cdots,m_k) = F_{ij}^d d_j + F_{ij}^h h_j + \sum_{\alpha=1}^{k} g_i^\alpha(\mathbf{x}) * m_\alpha(\mathbf{x}) = 0$，$(i = 1,\cdots,n)$

（3.5.21）

其中算子 F_{ij}^d， F_{ij}^h 和 $g_i^\alpha = \dfrac{\partial f_i}{\partial m_\alpha}$ 不含有模型参数和函数以及数据分量。

反问题的解：

$$\tilde{m}_\alpha(\mathbf{x}) = m_\alpha^0(\mathbf{x}) - C_\alpha^{m_0}(\mathbf{x},\mathbf{x}') * g_j^\alpha(\mathbf{x}') S_{jl}^{-1}\left\{F_{ln}^d d_n^0 + F_{ln}^h h_n^0 + \sum_{\beta=1}^{k} g_l^\beta(\mathbf{x}') * m_\beta^0(\mathbf{x}')\right\} \quad (3.5.22)$$

$$\tilde{h}_i = h_i^0 - C_{ik}^{h_0} F_{jk}^h S_{jl}^{-1}\left\{F_{ln}^d d_n^0 + F_{ln}^h h_n^0 + \sum_{\beta=1}^{k} g_l^\beta(\mathbf{x}') * m_\beta^0(\mathbf{x}')\right\} \quad (3.5.23)$$

$$\tilde{d}_i = d_i^0 - C_{ik}^{d_0} F_{jk}^d S_{jl}^{-1}\left\{F_{ln}^d d_n^0 + F_{ln}^h h_n^0 + \sum_{\beta=1}^{k} g_l^\beta(\mathbf{x}') * m_\beta^0(\mathbf{x}')\right\} \quad (3.5.24)$$

$$S_{ij} = F_{in}^d C_{nl}^{d_0} F_{jl}^d + F_{in}^h C_{nl}^{h_0} F_{jl}^h + \sum_{\alpha=1}^{k} g_i^\alpha(\mathbf{x}') * C_\alpha^{m_0}(\mathbf{x}',\mathbf{x}'') * g_j^\alpha(\mathbf{x}') \quad (3.5.25)$$

2)　数据可以用模型函数显式表示

正问题：

$$f(\mathbf{d},\mathbf{h},m_1(\mathbf{x}),\cdots,m_k(\mathbf{x})) = d_i - g_i(\mathbf{h},m_1(\mathbf{x}),\cdots,m_k(\mathbf{x})) = 0，\ (i=1,\cdots,n) \tag{3.5.26}$$

反问题的解：

$\dfrac{\partial f_i}{\partial m_j} = -\dfrac{\partial g_i}{\partial m_j}$， $F_{ij}^h = \dfrac{\partial f_i}{\partial h_j} = -\dfrac{\partial g_i}{\partial h_j}$， $F_{ij}^d = \dfrac{\partial f_i}{\partial d_j} = \delta_{ij}$，代入(3.5.17)－(3.5.20)可得

$$\begin{aligned} S_{ij} &= C_{ij}^{d_0} + \frac{\partial g_i}{\partial h_n} C_{nl}^{h_0} \frac{\partial g_j}{\partial h_l} \\ &+ \sum_{\alpha=1}^{k} \frac{\partial g_i(\mathbf{d},\mathbf{h},m_1(\mathbf{x}'),\cdots,m_k(\mathbf{x}'))}{\partial m_\alpha} * C_\alpha^{m_0}(\mathbf{x}',\mathbf{x}'') * \frac{\partial g_j(\mathbf{d},\mathbf{h},m_1(\mathbf{x}''),\cdots,m_k(\mathbf{x}''))}{\partial m_\alpha} \end{aligned} \tag{3.5.27}$$

$$\begin{aligned} h_i &= h_i^0 + C_{ik}^{h_0} \frac{\partial g_j}{\partial h_k} S_{jl}^{-1} \Big\{ \left(d_l^0 - d_l\right) + \frac{\partial g_l}{\partial h_n}\left(h_n - h_n^0\right) \\ &+ \sum_{\beta=1}^{k} \frac{\partial g_l(\mathbf{d},\mathbf{h},m_1(\mathbf{x}'),\cdots,m_k(\mathbf{x}'))}{\partial m_\beta} * \left(m_\beta(\mathbf{x}') - m_\beta^0(\mathbf{x}')\right) \Big\} \end{aligned} \tag{3.5.28}$$

$$\begin{aligned} m_\alpha(\mathbf{x}) &= m_\alpha^0(\mathbf{x}) + C_\alpha^{m_0}(\mathbf{x},\mathbf{x}') * \frac{\partial g_j(\mathbf{d},\mathbf{h},m_1(\mathbf{x}'),m_2(\mathbf{x}'),\cdots,m_k(\mathbf{x}'))}{\partial m_\alpha} S_{jl}^{-1} \\ &\left\{ \left(d_l^0 - d_l\right) + \frac{\partial g_l^h}{\partial h_n}\left(h_n - h_n^0\right) + \sum_{\beta=1}^{k} \frac{\partial g_l(\mathbf{d},\mathbf{h},m_1(\mathbf{x}'),\cdots,m_k(\mathbf{x}'))}{\partial m_\beta} * \left(m_\beta(\mathbf{x}') - m_\beta^0(\mathbf{x}')\right) \right\} \end{aligned} \tag{3.5.29}$$

$$\begin{aligned} d_i &= d_i^0 - C_{ij}^{d_0} S_{jl}^{-1} \\ &\left\{ \left(d_l^0 - d_l\right) + \frac{\partial g_l^h}{\partial h_n}\left(h_n - h_n^0\right) + \sum_{\beta=1}^{k} \frac{\partial g_l(\mathbf{d},\mathbf{h},m_1(\mathbf{x}'),\cdots,m_k(\mathbf{x}'))}{\partial m_\beta} * \left(m_\beta(\mathbf{x}') - m_\beta^0(\mathbf{x}')\right) \right\} \end{aligned} \tag{3.5.30}$$

对于非线性系统，（3.5.28）(3.5.30)是迭代解.

线性系统的解

正问题： $f_i(\mathbf{d},\mathbf{h},m_1,\cdots,m_k) = d_i - F_{ij}^h h_j - \sum_{\alpha=1}^{k} g_i^\alpha(\mathbf{x}) * m_\alpha(\mathbf{x}) = 0$ (3.5.31)

其中算子 F_{ij}^h， g_i^α 不含模型参数和模型函数以及数据分量。

反问题解：

$$S_{ij} = C_{ij}^{d_0} + F_{in}^h C_{nl}^{h_0} F_{jl}^h + \sum_{\alpha=1}^{k} g_i^\alpha(\mathbf{x}') * C_\alpha^{m_0}(\mathbf{x}',\mathbf{x}'') * g_j^\alpha(\mathbf{x}'') \qquad \text{（3.5.32}$$

(3.4.31)

$$\tilde{h}_i = h_i^0 + C_{ik}^{h_0} F_{jk}^h S_{jl}^{-1}\left\{d_l^0 - F_{ln}^h h_n^0 - \sum_{\beta=1}^{k} g_l^\beta(\mathbf{x}') * m_\beta^0(\mathbf{x}')\right\} \tag{3.5.33}$$

$$\tilde{m}_\alpha(\mathbf{x}) = m_\alpha^0(\mathbf{x}) + \left(C_\alpha^{m_0}(\mathbf{x},\mathbf{x}') * g_j^\alpha(\mathbf{x}')\right) S_{jl}^{-1}\left\{d_l^0 - F_{ln}^h h_n^0 - \sum_{\beta=1}^{k} g_l^\beta(\mathbf{x}') * m_\beta^0(\mathbf{x}')\right\} \tag{3.5.34}$$

$(\alpha = 1,\cdots,k)$

$$\begin{aligned}\tilde{d}_i &= d_i^0 - C_{ij}^{d_0} S_{jl}^{-1}\left\{d_l^0 - F_{ln}^h h_n^0 - \sum_{\beta=1}^{k} g_l^\beta(\mathbf{x}') * m_\beta^0(\mathbf{x}')\right\} \\ &= F_{ij}^h \tilde{h}_j + \sum_{\alpha=1}^{k} g_i^\alpha(\mathbf{x}') * \tilde{m}_\alpha(\mathbf{x}')\end{aligned} \tag{3.5.35}$$

3) 混合问题反演的分辨与误差

先讨论一组离散模型参数和一个模型函数的情况。设数据可以用模型函数显式表示。正问题为：$d_i - g_i(h_1,\cdots h_k, m(\mathbf{x})) = 0$，$(i = 1,\cdots,n)$

设先验模型协方差函数为$C^{m_0}(\mathbf{x},\mathbf{x})$，模型参数先验均值$h_i^0$的协方差矩阵为$C_{ij}^{h_0}$。模型函数及相关的量按照(3.4.2) －（3.4.10）按基函数展开表示后，参考（3.1.58），设矩阵$\mathbf{G}_1$的元素由$\dfrac{\partial g_i}{\partial p_j}$构成，矩阵$\mathbf{G}_2$的元素由$\dfrac{\partial g_i}{\partial h_j}$构成，则$\mathbf{G} = \begin{pmatrix}\mathbf{G}_1 & \mathbf{G}_2\end{pmatrix}$。由于先验模型函数相互独立，则 $\mathbf{C}_0^{mm} = \begin{pmatrix}\mathbf{C}^{p_0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}^{h_0}\end{pmatrix}$，令后验协方差矩阵为$\mathbf{C}^{mm} = \begin{pmatrix}\mathbf{C}^{pp} & \mathbf{C}^{ph} \\ \mathbf{C}^{hp} & \mathbf{C}^{hh}\end{pmatrix}$，根据（3.1.58）后验协方差矩阵可表示为

$$\begin{pmatrix}\mathbf{C}^{pp} & \mathbf{C}^{ph} \\ \mathbf{C}^{hp} & \mathbf{C}^{hh}\end{pmatrix} = \begin{pmatrix}\mathbf{C}^{p_0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}^{h_0}\end{pmatrix} - \begin{pmatrix}\mathbf{C}^{p_0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}^{h_0}\end{pmatrix}\begin{pmatrix}\mathbf{G}_1^T \\ \mathbf{G}_2^T\end{pmatrix}\mathbf{S}^{-1}\begin{pmatrix}\mathbf{G}_1 & \mathbf{G}_2\end{pmatrix}\begin{pmatrix}\mathbf{C}^{p_0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}^{h_0}\end{pmatrix}$$
$$== \begin{pmatrix}\mathbf{C}^{p_0} - \mathbf{C}^{p_0}\mathbf{G}_1^T\mathbf{S}^{-1}\mathbf{G}_1\mathbf{C}^{p_0} & -\mathbf{C}^{p_0}\mathbf{G}_1^T\mathbf{S}^{-1}\mathbf{G}_2\mathbf{C}^{h_0} \\ -\mathbf{C}^{h_0}\mathbf{G}_2^T\mathbf{S}^{-1}\mathbf{G}_1\mathbf{C}^{p_0} & \mathbf{C}^{h_0} - \mathbf{C}^{h_0}\mathbf{G}_2^T\mathbf{S}^{-1}\mathbf{G}_2\mathbf{C}^{h_0}\end{pmatrix}$$

$$\mathbf{S} = \mathbf{G}\mathbf{C}_0^{mm}\mathbf{G}^T + \mathbf{C}^{d_0} = \mathbf{G}_1\mathbf{C}^{p_0}\mathbf{G}_1^T + \mathbf{G}_2\mathbf{C}^{h_0}\mathbf{G}_2^T + \mathbf{C}^{d_0}$$

即

$$\mathbf{C}^{pp} = \mathbf{C}^{p_0} - \mathbf{C}^{p_0}\mathbf{G}_1^T\mathbf{S}^{-1}\mathbf{G}_1\mathbf{C}^{p_0}$$
$$\mathbf{C}^{hh} = \mathbf{C}^{h_0} - \mathbf{C}^{h_0}\mathbf{G}_2^T\mathbf{S}^{-1}\mathbf{G}_2\mathbf{C}^{h_0}$$
$$\mathbf{C}^{ph} = -\mathbf{C}^{p_0}\mathbf{G}_1^T\mathbf{S}^{-1}\mathbf{G}_2\mathbf{C}^{h_0}$$
$$\mathbf{C}^{hp} = -\mathbf{C}^{h_0}\mathbf{G}_2^T\mathbf{S}^{-1}\mathbf{G}_1\mathbf{C}^{p_0}$$

返回到连续模型，根据求(3.2.57)相同的步骤，就得到模型函数后验协方差及交叉协方差函数

$$S_{ij} = C_{ij}^{d_0} + \frac{\partial g_i}{\partial h_l} C_{ln}^{h_0} \frac{\partial g_j}{\partial h_n} + \frac{\partial g_i(\mathbf{h}, m(\mathbf{x}'))}{\partial m} * C^{m_0}(\mathbf{x}', \mathbf{x}'') * \frac{\partial g_j(\mathbf{h}, m(\mathbf{x}''))}{\partial m} \tag{3.5.36}$$

模型函数的后验协方差为

$$C^m(\mathbf{x}, \mathbf{x}') = C^{m_0}(\mathbf{x}, \mathbf{x}') - C^{m_0}(\mathbf{x}, \mathbf{x}'') * \frac{\partial g_i(\mathbf{h}, m(\mathbf{x}''))}{\partial m} S_{ij}^{-1} \frac{\partial g_j(\mathbf{h}, m(\mathbf{x}'''))}{\partial m} * C^{m_0}(\mathbf{x}''', \mathbf{x}') \tag{3.5.37}$$

离散模型参数的后验协方差矩阵为

$$C_{ij}^{hh} = C_{ij}^{h_0} - C_{ik}^{h_0} \frac{\partial g_l}{\partial h_k} S_{lm}^{-1} \frac{\partial g_m}{\partial h_n} C_{nj}^{h_0} \tag{3.5.38}$$

离散模型参数与模型函数交叉协方差函数为

$$C_i^{hm}(\mathbf{x}) = -C_{ij}^{h_0} \frac{\partial g_k}{\partial h_j} S_{kl}^{-1} \frac{\partial g_l(\mathbf{h}, m(\mathbf{x}'))}{\partial m} * C^{m_0}(\mathbf{x}', \mathbf{x}) \tag{3.5.39}$$

根据以上的分析很容易证明，当模型函数一共有 k 个时，正演条件为

$$d_i - g_i(h_1, \cdots h_k, m_1(\mathbf{x}), \cdots, m_k(\mathbf{x})) = 0$$

第 α 个模型函数的后验协方差函数为

$$C_{\alpha\alpha}^m(\mathbf{x}, \mathbf{x}') = C_\alpha^{m_0}(\mathbf{x}, \mathbf{x}')$$
$$- C_\alpha^{m_0}(\mathbf{x}, \mathbf{x}'') * \frac{\partial g_i(m_1(\mathbf{x}''), \cdots, m_k(\mathbf{x}''))}{\partial m_\alpha} S_{ij}^{-1} \frac{\partial g_j(m_1(\mathbf{x}'''), \cdots, m_k(\mathbf{x}'''))}{\partial m_\alpha} * C_\alpha^{m_0}(\mathbf{x}''', \mathbf{x}')$$

（3.5.40）

第 α 个模型函数与第 β 个模型函数的后验交叉协方差函数为

$$C_{\alpha\beta}^m(\mathbf{x}, \mathbf{x}') = -C_\alpha^{m_0}(\mathbf{x}, \mathbf{x}'') * \frac{\partial g_i(m_1(\mathbf{x}''), \cdots, m_k(\mathbf{x}''))}{\partial m_\alpha}$$

$$S_{ij}^{-1} \frac{\partial g_j(m_1(\mathbf{x}'''), \cdots, m_k(\mathbf{x}'''))}{\partial m_\beta} * C_\beta^{m_0}(\mathbf{x}''', \mathbf{x}') \tag{3.5.41}$$

离散模型参数的后验协方差矩阵为

$$C_{ij}^{hh} = C_{ij}^{h_0} - C_{ik}^{h_0} \frac{\partial g_l}{\partial h_k} S_{lm}^{-1} \frac{\partial g_m}{\partial h_n} C_{nj}^{h_0} \tag{3.5.42}$$

第i个离散模型参数与第α个模型函数交叉协方差函数为

$$C_{i\alpha}^{hm}(\mathbf{x}) = -C_{ij}^{h_0} \frac{\partial g_k}{\partial h_j} S_{kl}^{-1} \frac{\partial g_l(\mathbf{h}, m_1(\mathbf{x}'), \cdots, m_k(\mathbf{x}'))}{\partial m_\alpha} * C_\alpha^{m_0}(\mathbf{x}', \mathbf{x}) \tag{3.5.43}$$

$$\begin{aligned} S_{ij} &= C_{ij}^{d_0} + \frac{\partial g_i}{\partial h_l} C_{ln}^{h_0} \frac{\partial g_j}{\partial h_n} \\ &+ \sum_{\alpha=1}^{k} \frac{\partial g_i(m_1(\mathbf{x}'), \cdots, m_k(\mathbf{x}'))}{\partial m_\alpha} * C_\alpha^{m_0}(\mathbf{x}', \mathbf{x}'') * \frac{\partial g_j(m_1(\mathbf{x}''), \cdots, m_k(\mathbf{x}''))}{\partial m_\alpha} \end{aligned} \tag{3.5.44}$$

对于非线性问题，（3.4.36）－（3.4.38）中的模型函数取最后一次的迭代值。当$\mathbf{x} = \mathbf{x}'$时，$C_{\alpha\alpha}^{m}(\mathbf{x}, \mathbf{x})$表示第$\alpha$个模型函数在$\mathbf{x}$点的方差。当$\mathbf{x} \neq \mathbf{x}'$时，取定$\mathbf{x}$，$C_{\alpha\alpha}^{m}(\mathbf{x}, \mathbf{x}')$表示模型函数在$\mathbf{x}$点的分辨。由（3.4.37）可知，虽然假定了先验模型函数之间相互独立，但后验模型函数之间是有可能相关的。

几点说明

- 在连续模型函数的反演中，所有的偏导数都是Freshet导数，它们都是算子。算子的作用符“＊”作用到与算子有相同自变量的对象上。左作用与右作用的意义相同。
- 本讲义采用下脚标求和规则。所有公式中如果有相同的拉丁字母下脚标，表示求和。相同希腊字母的下脚标，不求和。

附录

A. 数学准备

符号规则：黑体小写拉丁字母代表矢量，黑体大写拉丁字母代表矩阵（非一维）。
带下角标的非黑体拉丁字母代表分量。

A.1 张量算符及下角标求和规则

下角标求和规则：下角标相同的两个拉丁字母代表求和。下角标为希腊字母不求和。

例：$a_i b_i = \sum_{\alpha=1}^{n} a_\alpha b_\alpha = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$，其中 n 是矢量 $\mathbf{a}$，$\mathbf{b}$ 的维数。

$$(\mathbf{AB})_{ij} = \sum_{\alpha=1} A_{i\alpha} B_{\alpha j} = A_{ik} B_{kj}$$

$$(\mathbf{A}^T \mathbf{a})_j = a_i A_{ij};$$

几个重要的运算符

a. δ_{ij} 算符

定义：$\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$

例：$\delta_{ij} a_j = a_i$，传递性。把角标 i 传给 j。

b. ε_{ijk} 算符

定义：
$$\varepsilon_{ijk}=\begin{cases}0 & ,\text{任意两个下角标相同}\\ 1 & ,ijk\text{取值}123,312,231\\ -1 & ,ijk\text{取值}321,132,213\end{cases}$$

例：$\varepsilon_{lmk}\varepsilon_{ijk}=\delta_{li}\delta_{mj}-\delta_{lj}\delta_{mi}$

c. $(\mathbf{a}\times\mathbf{b})_i=\varepsilon_{ijk}a_jb_k$ ，$\mathbf{a}\cdot\mathbf{b}=a_ib_i$ ，迪卡尔坐标 $i=1,3$

d. $(\nabla\times\mathbf{a})_i=\varepsilon_{ijk}\dfrac{\partial a_k}{\partial x_j}$，迪卡尔坐标 $i=1,3$

$\nabla\cdot\mathbf{a}=\dfrac{\partial a_i}{\partial x_i}$， 迪卡尔坐标 $i=1,3$

例题：证明 $\mathbf{u}\times(\mathbf{v}\times\mathbf{w})=(\mathbf{u}\cdot\mathbf{w})\mathbf{v}-(\mathbf{u}\cdot\mathbf{v})\mathbf{w}$

证：等式的左边$=\varepsilon_{ijk}u_j\varepsilon_{klm}v_lw_m$

$$=\varepsilon_{ijk}\varepsilon_{klm}u_jv_lw_m$$

$$=(\delta_{il}\delta_{jm}-\delta_{im}\delta_{jl})u_jv_lw_m$$

$$=u_mv_iw_m-u_lv_lw_i$$

$$=(\mathbf{u}\cdot\mathbf{w})v_i-(\mathbf{u}\cdot\mathbf{v})w_i=\text{右边}。$$ 证毕。

习题：（1）证明 $\nabla\times\nabla\times\vec{f}=\nabla(\nabla\cdot\vec{f})-\nabla^2\vec{f}$

（2）证明 $\nabla\times(\nabla\varphi)=0$，$\nabla\cdot(\nabla\times\vec{f})=0$

A2. 线性代数

a. 线性空间

- 加法运算

  在非空的集合 V 中，给定了一个对应规则，即一个算法，它使 V 中的任意两个元素 $\mathbf{a},\mathbf{b}$ 对应到 V 中的一个元素 $\mathbf{c}$，记为 $\mathbf{a}+\mathbf{b}=\mathbf{c}$。这算法遵守：

  （i）$\mathbf{a}+\mathbf{b}=\mathbf{b}+\mathbf{a}$

  （ii）$(\mathbf{a}+\mathbf{b})+\mathbf{x}=\mathbf{a}+(\mathbf{b}+\mathbf{x})$

  （iii）V 中存在元素 $\boldsymbol{\theta}$：是对 V 中的任意元素 $\mathbf{x}$，$\boldsymbol{\theta}+\mathbf{x}=\mathbf{x}$

  （iv）对于 V 中每一个 $\mathbf{x}$，存在 $(-\mathbf{x})\in V$，使 $\mathbf{x}+(-\mathbf{x})=\boldsymbol{\theta}$

- 数乘运算

在集合 V 中，给定一个对应规则，对任意取定的 $\mathbf{a}\in V$ ， $\lambda\in F$ (数域) ，对应 V 中的一个元素 $\mathbf{y}$ ，记为 $\lambda\mathbf{a}=\mathbf{y}$ ，这算法遵守：

（i）$\lambda(\mathbf{a}+\mathbf{b})=\lambda\mathbf{a}+\lambda\mathbf{b}$

（ii）$(\lambda+\mu)\mathbf{a}=\lambda\mathbf{a}+\mu\mathbf{a}$

（iii）$\lambda(\mu\mathbf{a})=(\lambda\mu)\mathbf{a}$

（iv）$1\mathbf{a}=\mathbf{a}$

- 线性空间

若取定数域 F，在非空集合 V 中，规定了加法运算 $\mathbf{a}+\mathbf{b}$ 和对域 F 中的数 λ 的数乘运算 $\lambda\mathbf{a}$ ，则 V 是域 F 上的线性空间或称向量空间。记为 $V(F)$，V 中的元素通常称为向量。

- 向量组的线性无关

给定 V 中的一组向量 $\mathbf{a}_1,\mathbf{a}_2,\cdots\mathbf{a}_m$ ，如果其中至少有一个，比如 $\mathbf{a}_1$ ，能表为组中其余向量的线性组合： $\mathbf{a}_1=\lambda_2\mathbf{a}_2+\lambda_3\mathbf{a}_3+\cdots+\lambda_m\mathbf{a}_m$ 。我们说向量组 $\mathbf{a}_1,\mathbf{a}_2,\cdots\mathbf{a}_m$ 是线性相关的。

定义：对于向量组 $\mathbf{a}_1,\mathbf{a}_2,\cdots\mathbf{a}_m$ ，如果 $\mathbf{a}_1\lambda_1+\lambda_2\mathbf{a}_2+\lambda_3\mathbf{a}_3+\cdots+\lambda_m\mathbf{a}_m=\boldsymbol{\theta}$ ，仅当 $\lambda_1=\lambda_2=\cdots=\lambda_m=0$ 时成立，则称这向量组是线性无关的。

- 基和坐标

定义：线性空间 V 中，如果他的极大线性无关组所含向量的个数是 n, 就称 V 是 n 维线性空间。V 中任意一个极大线性无关组 $\{\mathbf{a}_1,\mathbf{a}_2,\cdots\mathbf{a}_n\}$ 称为 V 的一组基。V 中的任意一个向量 $\mathbf{a}$ 都可以用基来线性表示

$$\mathbf{a}=a_1\mathbf{a}_1+a_2\mathbf{a}_2+\cdots+a_n\mathbf{a}_n$$

系数组 $(a_1,a_2,\cdots a_n)$ 称为 $\mathbf{a}$ 关于基 $\{\mathbf{a}_1,\mathbf{a}_2,\cdots\mathbf{a}_n\}$ 的坐标。

只要线性空间 V 中含有一个非零向量 $\mathbf{a}$ ，$\{\mathbf{a}\}$ 就是一个线性无关组，空间的维数就 ≥ 1；当线性无关组含有无穷多个向量时，空间就是 ∞ 维的；当 V 中仅含有零向量时，自然就称他为 0 维空间。

b. 线性变换

- $V_n(F)$ 中的变换

$V_n(F)$ 中的变换 $\Re$ 定义为： $\mathbf{y}=\Re\mathbf{x}\in V_n(F)$ ， $\mathbf{x}\in V_n(F)$ 。

- 线性变换

定义：若$\Re$是$V(F)$的变换，他把 V 中的每一个向量$\mathbf{x}$变换到$\Re\mathbf{x}\in V$，并满足

$$\Re(\mathbf{x}+\mathbf{y})=\Re\mathbf{x}+\Re\mathbf{y}\text{；}\Re(\lambda\mathbf{x})=\lambda(\Re\mathbf{x})\text{，}\quad\lambda\in F$$

则称$\Re$是V的线性变换或称线性算子。

- 线性变换的坐标表示