The background of the slide is a painting. On the left, a woman in a light-colored dress and a wide-brimmed hat sits on a wooden bench, holding a small child. On the right, another woman in a dark blue long-sleeved top and a long brown skirt stands with her back to the viewer, looking at a bush of red and white flowers. She is holding a woven basket. The scene is set in a garden with a large tree on the left and a building with a thatched roof on the right.

# 反问题笔记

Inverse Problem Notes

讲学笔记

RESEARCH NOTE, AT UNIVERSITY OF HOUSTON

[GITHUB.COM/CAITMAGI/???](https://github.com/caitmagi/???)

This is the personal note of learning inverse problem theory, written by *Yuchen Jin*. Anyone who read this book could share it freely, but any activities of recompilation, modification or reproduction are not permitted, according to the authority of the editor.

*First release, February 9, 2018*






## 目 录

<b>1</b>	<b>普适离散反问题 .....</b>	<b>5</b>
1.1	<b>模型空间与数据空间 .....</b>	<b>5</b>
1.1.1	模型空间 .....	6
1.1.2	数据空间 .....	7
1.1.3	合并流形 .....	7





## 1. 普适离散反问题

反问题 (*inverse problem*) 与正问题 (*forward problem*, *modelization problem*, *simulation problem*) 相对, 是基于已经存在的观测结果, 逆推出产生如此观测结果的模型。一个最大的不同是, 正问题里确定模型时, 观测结果有唯一解; 反问题里确定观测结果时, 模型没有唯一解。

故而, 为了得到有效的解、或解集 (*collection of models*), 往往须要对问题施加约束条件, 这种约束条件可以看成是对模型的一种先验信息, 它定义了解集可能存在的概率分布 (因而我们可以探索到最大概率的解)。

当数据有较高的采样频率时, 可以描述一个离散反问题 (*discrete inverse problem*) 来代替原问题。离散反问题从数学上更容易解, 但它不能完全等同于原问题, 因为离散化的过程中引入了一些简化, 造成一些信息因为离散化被隐藏了起来。

本章的核心是参数集 (*parameter set*) 上的“状态信息 (*state of information*)”, 一般被描述为前述的先验信息/概率密度。这种方法可以用在完全非线性问题里, 但计算开销甚大。且它的解一般少于传统方法得到的解。

### 1.1 模型空间与数据空间

研究一个物理系统  $\mathcal{G}$  一般分为三步:

- (1) **系统参数化** (*parameterization of system*): 确定一个最小的参数集, 使得通过这些参数可以完全描述物理系统;
- (2) **正向模型** (*forward modeling*): 预测并给定一些参数, 通过物理定律, 建立参数  $\rightarrow$  观测结果之间的关系;
- (3) **反向模型** (*inverse modeling*): 使用这些预测的参数, 推测出实际模型参数的解。

可见, 前两步是归纳 (*inductive*), 最后一步是演绎 (*deductive*), 前两步的物理建模可能更困难; 至于最后一步, 有成熟、有效的数学方法来解。

首先让我们考量系统参数化的问题:

■ 例 1.1 考虑一个弹性系统  $\mathcal{G}$ , 对于空间上任意的点  $\mathbf{x}$ , 有压力  $\sigma^{ij}(\mathbf{x})$ , 拉力  $\varepsilon^{ij}(\mathbf{x})$ , 则可以定义弹性硬度张量  $c_{kl}^{ij}(\mathbf{x})$  为:



$$\sigma^{ij}(\mathbf{x}) = c_{kl}^{ij}(\mathbf{x}) \varepsilon^{kl}(\mathbf{x}). \quad (1.1)$$

同时还可以定义弹性柔度张量  $s_{kl}^{ij}(\mathbf{x})$  为:

$$\varepsilon^{ij}(\mathbf{x}) = s_{kl}^{ij}(\mathbf{x}) \sigma^{kl}(\mathbf{x}). \quad (1.2)$$

这里, 张量  $\mathbf{s}$  是张量  $\mathbf{c}$  的逆, 亦即  $c_{kl}^{ij} s_{mn}^{kl} = \delta_m^i \delta_n^j$ 。故而选择这两个参数中的任何一个都是**完全等价** (*completely equivalent*) 的。■

选择任意一组符合上述要求的特定参数集就是**系统参数化**, 对于两个等价的参数集, 其参数之间可以一一映射。

### 1.1.1 模型空间

与系统参数化无关, 我们可以引入一个**抽象空间**, 或称为**流形** (*manifold*), 用其中的每一个**点**代表一个模型  $\mathbf{m}$ , 于是这样的流形记为**模型空间** (*model space*)  $\mathfrak{M} = \{\mathbf{m}_1, \mathbf{m}_2, \dots\}$ 。

一个**模型空间**可以被不同的 (等价) 系统参数描述。例如, 设若  $\mathfrak{M}$  的参数有两维, 记为  $\{m^1, m^2\}$ , 则  $m^1$  和  $m^2$  可以选用不同的参数进行测量。由此定义的**模型**, 或记为**点** $\mathbf{m}$ , 一般来说不可以看成是一个 (线性空间内的) 向量, 因为其坐标轴可能是非线性的。即使坐标轴都是线性的, 也一般不是笛卡尔坐标系。

■ **例 1.2** 考虑一个**模型空间** $\mathfrak{M}$  已经确定了系统参数为  $\{\kappa, \mu\}$  (分别为体积模量和剪切模量), 可以定义两个模型  $\mathbf{m}_1, \mathbf{m}_2$  之间的距离为

$$d = \sqrt{\left(\log \frac{\kappa_2}{\kappa_1}\right)^2 + \left(\log \frac{\mu_2}{\mu_1}\right)^2} \quad (1.3)$$

若以对数化的系统参数  $\{\kappa^* = \log \kappa, \mu^* = \log \mu\}$  来代替(1.3)的参数, 有:

$$d = \sqrt{(\kappa_2^* - \kappa_1^*)^2 + (\mu_2^* - \mu_1^*)^2} \quad (1.4)$$

可见, (1.4) 是的参数是笛卡尔坐标系, 但(1.3)的不是。■

**性质 1.1.1 – 离散性.** 须知的是, 一个实际物理模型, 其**参数系** (*set of quantities*) (亦即坐标轴) 往往是连续的, 记为  $m(\mathbf{x})$ ; 但本章考虑的模型是经过抽样、离散化的模型, 其**参数系**为  $m^\alpha = m^\alpha(\mathbf{x})$ 。坐标轴的离散性满足以下条件:

- (1) 参数系由选取的系统参数决定, 一个模型可以选用不同的参数系;
- (2) 参数系维数  $\alpha$  有限, 本章不考虑无限的情况;
- (3) 参数系的抽样应该足够多, 能不失真地反映出原始模型的特征;
- (4) 参数系的值本身可以取连续也可以取离散, 为了方便表达, 可以默认认为取值总是连续的, 因为离散的取值可以用**狄拉克** (*Dirac*) 函数表达。■

**性质 1.1.2 – 线性.** 设任意两个同属一个模型空间  $\mathfrak{M}$ , 使用同一系统参数  $m^\alpha$  的模型  $\mathbf{m}_1, \mathbf{m}_2$ , 其任意第  $i$  维的值均满足:

$$(\mathbf{m}_1 + \mathbf{m}_2)^\alpha = (\mathbf{m}_1)^\alpha + (\mathbf{m}_2)^\alpha \quad (1.5)$$

$$(\lambda \mathbf{m})^\alpha = \lambda (\mathbf{m})^\alpha \quad (1.6)$$

这样的  $\mathfrak{M}$  称为**线性模型空间** (*linear model space*), 记为  $\mathbb{M}$

例 1.2使用的对数化系统参数  $\{\kappa^*, \mu^*\}$  是可以构成一个**线性模型空间**。因此, 此例中的距离  $d$  相当于该线性空间的**Euclidean 范数**。

为了处理一个反问题, 设模型空间  $\mathfrak{M}$  为全集, 其任意一个子集记为  $\mathcal{A}$ , 则可以令全集概率  $P(\mathfrak{M}) = 1$ , 且其子集概率  $P(\mathcal{A}) \in [0, 1]$ 。此处定义的概率, 是通过一种特定的关联关系定义的, 这种关系不受到空间是否是线性的影响。并且, 能相应地根据该概率求出每个模型的概率密度分布。

### 1.1.2 数据空间

对于一个模型的观测结果, 由于观测仪器的不同也可能产生诸多不同的结果, 设全集  $\mathfrak{D}$  为所有可观测仪器产生的观测结果的集合, 那么任何一组来自一个仪器的测量结果都可以看成它的元素  $\mathbf{d}$ ; 与  $\mathfrak{M}$  相似,  $\mathfrak{D}$  同样也可以看成是一个**流形**,  $\mathbf{d}$  可以看成是流形中的一个**点**。

与性质 1.1.1和性质 1.1.2相似, 我们讨论的数据空间也有类似的性质。离散性此处从略, 在此我们介绍线性:

**性质 1.1.3 – 线性.** 在一个线性数据空间  $\mathbb{D}$  内, 任意两组测量结果 (来自两个仪器)  $\mathbf{d}_1, \mathbf{d}_2$ , 其任意第  $i$  维的值均满足:

$$(\mathbf{d}_1 + \mathbf{d}_2)^\alpha = (\mathbf{d}_1)^\alpha + (\mathbf{d}_2)^\alpha \quad (1.7)$$

$$(\lambda \mathbf{d})^\alpha = \lambda (\mathbf{d})^\alpha \quad (1.8)$$

### 1.1.3 合并流形

有时候, 一个物理系统的输入和输出参数可以清楚地分开, 这时使用两个**流形**  $\mathfrak{M}, \mathfrak{D}$  来描述并无问题。但某些情况下 (例如有反馈系统), 不能做到这一点。因此有时我们也用一个合并的**流形**  $\mathfrak{X}$  来表示整个**参数空间** (*parameter space*)。这种情况下, 其元素 (或**点**)  $\mathbf{x}$  被称为一组参数。