SVD 展开方法详解（Höcker–Kartvelishvili Unfolding）

1. 基本设定

待测物理量为真实的刚度谱 $f(R_{\text{true}})$，探测器实测谱为 $g(R_{\text{meas}})$。将连续谱离散化到 n 个测量 bin 和 m 个真实 bin，记 f_j 为真实第 j 个 bin 的事例数，g_i 为测量第 i 个 bin 的事例数。探测器响应用**迁移矩阵** A_{ij} 表示：A_{ij} 是真实处于 bin j 的事例被测量到 bin i 的概率，满足 $\sum_i A_{ij} = 1$。

于是有线性模型

$$g_i = \sum_{j=1}^{m} A_{ij} f_j + \varepsilon_i$$

或写为矩阵形式

$$\boldsymbol{g} = \mathbf{A}\,\boldsymbol{f} + \boldsymbol{\varepsilon}$$

其中 $\boldsymbol{\varepsilon}$ 为测量统计噪声（通常假定服从高斯分布，协方差矩阵 $\mathbf{V}$ 可估）。

目标：已知测量值 $\boldsymbol{g}$、迁移矩阵 $\mathbf{A}$（由蒙特卡罗模拟确定），估计真实谱 $\boldsymbol{f}$。

---

2. 直接反演的病态性

如果用最小二乘法直接最小化 $\chi^2 = (\boldsymbol{g} - \mathbf{A}\boldsymbol{f})^\top \mathbf{V}^{-1} (\boldsymbol{g} - \mathbf{A}\boldsymbol{f})$，正规方程为

$$\mathbf{A}^\top \mathbf{V}^{-1} \mathbf{A}\,\boldsymbol{f} = \mathbf{A}^\top \mathbf{V}^{-1} \boldsymbol{g}$$

解得

$$\boldsymbol{f} = (\mathbf{A}^\top \mathbf{V}^{-1} \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{V}^{-1} \boldsymbol{g}$$

但迁移矩阵 $\mathbf{A}$ 是病态的：相邻 bin 的迁移概率高度相关，小奇异值对应的模式对噪声极其敏感。直接求逆会产生剧烈振荡的解，甚至出现负值事例数。

SVD 方法正是通过**截断小奇异值**来正则化，牺牲部分分辨率以换取解的稳定性。

---

3. SVD 分解与正则化

3.1 规范化矩阵

为简化，引入**规范化协方差矩阵**，通常使用对数似然拟合或平方根矩阵。经典做法是定义

$$\mathbf{C} = \mathbf{V}^{-1/2}\mathbf{A} \qquad \text{和} \qquad \boldsymbol{y} = \mathbf{V}^{-1/2}\boldsymbol{g}$$

其中 $\mathbf{V}^{-1/2}$ 是 $\mathbf{V}^{-1}$ 的对称平方根。则将问题转化为普通最小二乘：

$$\chi^2 = \|\boldsymbol{y} - \mathbf{C}\boldsymbol{f}\|^2$$

正规方程变成 $\mathbf{C}^\top\mathbf{C}\,\boldsymbol{f} = \mathbf{C}^\top\boldsymbol{y}$。

3.2 奇异值分解

对矩阵 $\mathbf{C}$（尺寸 $n \times m$，通常 $n \geq m$）进行 SVD：

$$\mathbf{C} = \mathbf{U}\,\boldsymbol{\Sigma}\,\mathbf{W}^\top$$

其中：

- $\mathbf{U}$ 是 $n \times n$ 正交矩阵，列向量 $\{\boldsymbol{u}_k\}$ 是左奇异向量；
- $\mathbf{W}$ 是 $m \times m$ 正交矩阵，列向量 $\{\boldsymbol{w}_k\}$ 是右奇异向量；
- $\boldsymbol{\Sigma}$ 是 $n \times m$ 对角矩阵，对角元 $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_m \geq 0$ 为奇异值。

3.3 正则化解的形式

利用 SVD，正规方程 $\mathbf{C}^\top\mathbf{C}\,\boldsymbol{f} = \mathbf{C}^\top\boldsymbol{y}$ 的解可表为

$$\boldsymbol{f}_{\text{reg}} = \sum_{k=1}^{m} d_k\,\boldsymbol{w}_k$$

其中系数

$$d_k = \frac{\boldsymbol{u}_k^\top\boldsymbol{y}}{\sigma_k}$$

如果保留所有奇异值，这就是未经正则化的普通最小二乘解。由于小 σ_k 对应的项对噪声极其敏感（分母很小，放大了 $\boldsymbol{y}$ 中的统计涨落），SVD 正则化通过引入**衰减因子** ϕ_k 来抑制这些项：

$$\boldsymbol{f}_{\text{reg}} = \sum_{k=1}^{m} \phi_k\,\frac{\boldsymbol{u}_k^\top\boldsymbol{y}}{\sigma_k}\,\boldsymbol{w}_k$$

最简单的选择是硬截断：取某个正整数 $k_{\max} < m$，令

$$\phi_k = \begin{cases} 1, & k \leq k_{\max} \\ 0, & k > k_{\max} \end{cases}$$

更平滑的衰减函数（如 Tikhonov 型 $\phi_k = \frac{\sigma_k^2}{\sigma_k^2+\tau^2}$）也可引入。

---

4. 正则化参数的选择

硬截断时，截断阶数 $k_{\max}$（或 Tikhonov 参数 τ）决定了平滑程度：

- $k_{\max}$ 过大：保留太多噪声，解出现剧烈振荡；

- $k_{\max}$ 过小：解过于平滑，丢失真实谱的细节（偏差增大）。

选择方法：

- **与蒙特卡罗真值对比**：用已知真实谱的模拟数据测试不同 $k_{\max}$，选择使偏差和方差综合最优者；
- **L-曲线法**：绘制解的范数 $\|\boldsymbol{f}_{\text{reg}}\|^2$ 与残差范数 $\|\boldsymbol{y} - \mathbf{C}\boldsymbol{f}_{\text{reg}}\|^2$ 的双对数曲线，拐点对应最佳平衡点；
- **交叉验证**：在真实数据中谨慎使用，容易因 bin 间相关性产生误导。

---

5. 误差传播与协方差

SVD 展开的解是 $\boldsymbol{g}$ 的线性函数，因此解的后验协方差矩阵可直接计算：

$$\text{Cov}(\boldsymbol{f}_{\text{reg}}) = \sum_{k=1}^{m} \phi_k^2 \frac{\boldsymbol{w}_k \boldsymbol{w}_k^\top}{\sigma_k^2}$$

这反映了由于数据统计误差引起的解的统计不确定性。注意截断会同时引入偏差，完整的误差评估需同时考虑统计误差和正则化偏差。

---

6. 优缺点与使用注意

优点：

- 数学框架清晰，基于线性代数，实现简单；
- 正则化强度通过截断阶数直观控制；
- 适用于测量 bin 数不小于真实 bin 数的一般情况。

缺点：

- 假设测量误差为高斯分布且协方差矩阵已知，对低统计量或非高斯情形不稳健；
- 需要充分的蒙特卡罗模拟以构建准确的迁移矩阵和协方差矩阵；
- 硬截断会在 bin 之间产生与正则化阶数相关的相关性结构，需要谨慎诠释；
- 对小信号分量（在某些 bin 统计不足）可能仍有偏。

在径迹探测器刚度谱 unfolding 中，SVD 方法常与贝叶斯迭代方法对比。当探测分辨率造成的 bin 迁移主要限制在相邻几个 bin 时，SVD 正则化能有效消除统计振荡，并保持边缘的形状。实际使用时，一般会与多种方法交叉验证，确保结果不依赖于特定正则化策略。

---

7. 总结

SVD 展开通过在奇异值空间中截断小奇异值，滤除噪声主导的模式，使得线性反演问题从病态变为适定。其核心是**用偏差换取方差减小**，在提高谱分辨率与保证统计稳定性之间取得平衡。正确应用的关键在于选择合理的正则化参数，并对解的不确定性（包括统计误差和系统性偏差）进行全面估计。