



一个最优化问题的简化

Uwni ✉ satanyalin@gmail.com

目录

0.1 整理目標函數 2

我的研究过程中, 遇到了如下的优化问题. 现在考虑下面这种情形, 我有任意的 n 阶幺正矩阵 $U \in U(n)$ 对应了可实现的物理构造, 然后有 n 阶输入方阵为 $X, n = 2^m, X$ 的前 m 行是 m bit 的 BPSK 信号的所有组合, 数学上就是 Rademacher 空间, 后 $n - m$ 行为每行都相同的 $n - m$ 个自由未定参数. 也就是说 X 的每一列代表了一种 BPSK 编码的输入状态, 即 Rademacher 空间 $\{\pm 1\}^m$ 中的一个元素, 如若 $m = 3$ 的时候的话, X 可写成

$$\begin{bmatrix} 1 & 1 & \cdots & -1 & -1 \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & -1 & \cdots & 1 & -1 \\ a_1 & a_1 & \cdots & a_1 & a_1 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n-m} & a_{n-m} & \cdots & a_{n-m} & a_{n-m} \end{bmatrix}$$

每一行代表了一个输入信号通道.

写成分块矩阵为

$$X(a) = \begin{bmatrix} B \\ A \end{bmatrix}$$

其中 B 为 m 行的 BPSK 信号矩阵, 其列枚举了所有的 m bit BPSK 信号组合, A 为 $n - m$ 行的自由参数矩阵. 其各列皆为 $n - m$ 元向量 a . 即 $A = a \mathbf{1}_n^T \in \mathbf{C}^{(n-m) \times n}$. 考虑输入信号的每一种状态, 即 X 的每一列 $x_j = (b_j; a)^T$, 其中 b_j 为 B 的第 j 列. 现在考虑功率矩阵 P 为输出 UX 各元素的功率, 即有

$$p_{ij} = |(UX)_{ij}|^2 = |\langle u_i, x_j \rangle|^2, \quad u_i \in \mathbf{C}^{1 \times n}, x_j \in \mathbf{C}^{n \times 1}$$

并对 P 的每一列向量归一化后成为 P_{norm} . 我现在要这个物理器件在对以 a 为额外输入参数时, 成为一个译码器 (decoder) 的功能. 数学上, 即对幺正矩阵 U 和向量 a 对 $\|P_{\text{norm}} - I\|_F^2$ 求最小值优化, 其中 I 为同阶的单位阵. 也就是说要输出尽可能为 one-hot 的分布. 通过数值计算我发现所有的局部最优值都是全局最优 (即有多个局部最优但结果都一样), 现在要通过理论分析来证明这一点. . (over-parametrization)

0.1 整理目標函數

现在有几个重要的事实 显然有一点对于 $\mathbf{X}$ 的每一列其功率都是一样的, 即对于每一列 $\mathbf{x}_j = (b_j; \mathbf{a})^T$ 有

$$L := \|\mathbf{x}_j\|^2 = \|\mathbf{b}_j\|^2 + \|\mathbf{a}\|^2 = m + \|\mathbf{a}\|^2$$

即每一列的功率都同为 L . 那么优化问题自然便成了

$$\begin{aligned} \min_{\substack{\mathbf{U} \in U(n) \\ \mathbf{a} \in \mathbb{C}^{n-m}}} \|\mathbf{P} - L\mathbf{I}\|_F^2 &= \sum_{i=1}^n \sum_{j=1}^n (p_{ij} - L\delta_{ij})^2 \\ &= \sum_{i,j} |p_{ij}|^2 - 2L \sum_{i=1}^n \sum_{j=1}^n p_{ij}\delta_{ij} + nL^2 \\ &= \sum_{i,j} |p_{ij}|^2 - 2L \sum_{i=1}^n p_{ii} + nL^2 \end{aligned}$$

把 $\mathbf{U}$ 分块, 第 i 行 $\mathbf{u}_i \in \mathbb{C}^{1 \times n}$ 拆成

$$\mathbf{u}_i = [\mathbf{v}_i \ \mathbf{w}_i]$$

其中 $\mathbf{v}_i \in \mathbb{C}^{1 \times m}, \mathbf{w}_i \in \mathbb{C}^{1 \times (n-m)}$. 因为 $\mathbf{U}$ 么正, $\|\mathbf{v}_i\|^2 + \|\mathbf{w}_i\|^2 = 1$ 而 i 行, j 列, 所對應的输出元素為

$$y_{ij} = \langle \mathbf{u}_i, \mathbf{x}_j \rangle = \langle \mathbf{v}_i, \mathbf{b}_j \rangle + \langle \mathbf{w}_i, \mathbf{a} \rangle$$

觀察到 $\sum_j \mathbf{b}_j = 0, \sum_j \mathbf{b}_j \mathbf{b}_j^H = n\mathbf{I}$, 定義 $r_i := \|\mathbf{v}_i\|^2$, $\mathbf{r} := (r_i)$, 注意到 $0 \leq r_i \leq 1$

$$\sum_{i=1}^n r_i = \sum_{i=1}^n \|\mathbf{v}_i\|^2 = m$$

因為 $\mathbf{W}$ 为 $\mathbf{U}$ 的右 $(n-m)$ 列块, 么正矩阵的列正交归一化意味着 $\mathbf{W}^H \mathbf{W} = \mathbf{I}_{n-m}$, 所以定义

$$\sum_{i=1}^n |\langle \mathbf{w}_i, \mathbf{a} \rangle|^2 = \|\mathbf{W}\mathbf{a}\|^2 = \mathbf{a}^H \mathbf{W}^H \mathbf{W} \mathbf{a} = \|\mathbf{a}\|^2$$

對於固定的输出端口 i 在各个状态 j 下的平均功率有

$$\frac{1}{n} \sum_{j=1}^n p_{ij} = \frac{1}{n} \sum_{j=1}^n |y_{ij}|^2 = \frac{1}{n} \left(\sum_{j=1}^n |\langle \mathbf{v}_i, \mathbf{b}_j \rangle|^2 + |\langle \mathbf{w}_i, \mathbf{a} \rangle|^2 \right) =$$

讨论: 但是是不是在置换下其实是唯一的? 对呀, 这玩意一看就有对称性吧. 是不是某个更小的商群下, 问题是凸的. 我怀疑. permutation 和 phase rotation 之下最优解还唯一么, 需要验证. 这个 loss 展开大概是什么, 四阶矩么? 这样带进去, 假设 $\mathbf{Y} = \mathbf{U}\mathbf{X}$. 就是一个四次方. 我的猜想是, 商去某些自由度之后, 这个问题可能是凸的. 能否证明目标函数存在巨大的对称群? 把目标函数在群下的作用目标不变证明一下. 能否 check 一下你最后算出来的那些不同的 minima, 对应的 $\mathbf{U}\mathbf{X}$ 长啥样. 对生成元逐个证明, 在这些对称都扔掉, 取商空间之后呢, 是否还有局部极小值. 在数学上可以有. 光滑紧致流形上的全局极小点正交于对称轨道方向一般是正定的, 对称群的零模方向被 quotient 之后就剩下孤立的点了, $\mathbf{j}$ 所有切向量的二阶导数都大于 0. $\mathbf{J}(\mathbf{U}, \mathbf{a}) = 0$ 这样是吧.