具有连续权重的感知机在二分类任务中的泛化误差

李宇豪*‡　王誉晨†‡

日期：2023 年 6 月 5 日§

摘　要

我们在老师学生模型和贝叶斯最优框架中分析了具有连续权重的感知机在二分类任务中的泛化误差。我们首先基于信念传播方程（BP）推导了广义消息传递方程（GAMP），通过迭代得到不同数据量密度时的泛化误差；然后基于消息传递方程推导出了状态演化方程（SE），通过迭代得到了泛化误差随数据量密度变化的理论曲线，与使用广义消息传递方程得到的结果吻合；最后我们使用复本方法进行计算分析，验证了状态演化方程的结果，并且分析了在数据量密度趋于无穷大时泛化误差的渐近行为。

目录

*liyh536@mail2.sysu.edu.cn

†wangych256@mail2.sysu.edu.cn

‡李宇豪主要撰写了 Sec 1、2、4，完成了 GAMP 方程的推导和 Replica 计算，
王誉晨主要撰写了 Sec 3、5，完成了 SE 方程的推导和 GAMP 算法的编程。

§2023 年 11 月 6 日更新了复本部分的计算细节

1 Model setting and main results

1.1 模型设定

在老师-学生模型中考虑一个 $N-1$ 的二分类感知机，设输入数据为 $\mathbf{X}$，老师的权重 $\mathbf{w}^*$ 是 ground truth，并且满足高斯先验 $P_{\mathbf{w}^*} \sim \mathcal{N}(0,1)$，老师通过

$$y = \mathrm{sign}\left(\frac{1}{\sqrt{n}}\mathbf{X}\mathbf{w}^\star\right) \tag{1}$$

生成标签；学生根据数据学习老师的权重 $\hat{\mathbf{w}}$，即 $P(\mathbf{w}|\mathbf{y},\mathbf{X})$。在贝叶斯最优框架中有

$$P(\mathbf{w}|\mathbf{y},\mathbf{X}) = \frac{P(\mathbf{y}|\mathbf{X},\mathbf{w})P(\mathbf{w})}{P(\mathbf{y},\mathbf{X})} = \frac{1}{\mathcal{Z}}P(\mathbf{y}|\mathbf{X},\mathbf{w})P(\mathbf{w}) \tag{2}$$

其中，

$$P(\mathbf{y}|\mathbf{X},\mathbf{w}) = \mathbb{1}\left[y = \mathrm{sign}\left(\frac{1}{\sqrt{n}}\mathbf{X}\mathbf{w}\right)\right] \tag{3}$$

我们共使用 d 套数据进行训练，定义数据量密度 $\alpha = d/n$。

二分类任务的泛化误差（即预测标签和真实标签不匹配的概率）定义为

$$\varepsilon_{\text{gen}} = \lim_{n\to\infty}\mathbb{E}_{\mathbf{y},\mathbf{X}}\mathbb{1}[\mathbf{y} \neq \hat{y}(\hat{\mathbf{w}}(\alpha);\mathbf{x})] = \mathbb{E}_{\mathbf{y},\mathbf{X}}\left[\Theta(-\mathbf{y}\hat{\mathbf{y}})\right] = \mathbb{E}_{\mathbf{y},\mathbf{X}}\left[\Theta(-\mathbf{z}\hat{\mathbf{z}})\right] \tag{4}$$

其中 $\mathbf{z} = \mathbf{X}\mathbf{w}^\star/\sqrt{n}$，$\hat{\mathbf{z}} = \mathbf{X}\hat{\mathbf{w}}/\sqrt{n}$。注意到向量 $(\mathbf{z},\hat{\mathbf{z}})$ 在对所有可能的真实权重 $\mathbf{w}^\star$ 和输入数据 $\mathbf{X}$ 取平均后满足高斯分布 $\mathcal{N}(0,\sigma)$，其中协方差矩阵

$$\sigma = \lim_{n\to\infty}\mathbb{E}_{\mathbf{w}^*,X}\frac{1}{n}\begin{bmatrix}\mathbf{w}^{*\top}\mathbf{w}^* & \mathbf{w}^{*\top}\hat{\mathbf{w}} \\ \mathbf{w}^{*\top}\hat{\mathbf{w}} & \hat{\mathbf{w}}^\top\hat{\mathbf{w}}\end{bmatrix} \equiv \begin{bmatrix}\sigma_{\mathbf{w}^*} & \sigma_{\mathbf{w}^*\hat{\mathbf{w}}} \\ \sigma_{\mathbf{w}^*\hat{\mathbf{w}}} & \sigma_{\hat{\mathbf{w}}}\end{bmatrix} \tag{5}$$

将 $\mathbf{z}$ 和 $\hat{\mathbf{z}}$ 用标准高斯分布重参数化为

$$\begin{aligned}\hat{\mathbf{z}} &= \sqrt{\sigma_{\hat{\mathbf{w}}}}\mathbf{x}_1 \\ \mathbf{z} &= \sqrt{\frac{\sigma_{\mathbf{w}^*\hat{\mathbf{w}}}}{\sigma_{\hat{\mathbf{w}}}}}\mathbf{x}_1 + \sqrt{\sigma^2_{\mathrm{w}^*} - \frac{\sigma_{\mathbf{w}^*\hat{\mathbf{w}}}}{\sigma_{\hat{\mathbf{w}}}}}\mathbf{x}_2\end{aligned} \tag{6}$$

其中 $\mathbf{x}_1,\mathbf{x}_2 \sim \mathcal{N}(0,1)$。因此泛化误差可以重写为

$$\begin{aligned}\varepsilon_{\text{gen}} &= \mathbb{E}_{\mathbf{y},\mathbf{X}}\left[\Theta(-\mathbf{z}\hat{\mathbf{z}})\right] \\ &= \int D\mathbf{x}_1\int D\mathbf{x}_2\Theta\left(-\sqrt{\sigma_{\hat{\mathbf{w}}}}\mathbf{x}_1\left(\sqrt{\frac{\sigma^2_{\mathbf{w}^*\hat{\mathbf{w}}}}{\sigma_{\hat{\mathbf{w}}}}}\mathbf{x}_1 + \sqrt{\sigma^2_{\mathbf{w}^*} - \frac{\sigma^2_{\mathbf{w}^*\hat{\mathbf{w}}}}{\sigma_{\hat{\mathbf{w}}}}}\mathbf{x}_2\right)\right) \\ &= \int D\mathbf{x}_1\int D\mathbf{x}_2\Theta\left(-\mathbf{x}_1\right)\Theta\left(\sqrt{\frac{\sigma^2_{\mathbf{w}^*\hat{\mathbf{w}}}}{\sigma_{\hat{\mathbf{w}}}}}\mathbf{x}_1 + \sqrt{\sigma^2_{\mathbf{w}^*} - \frac{\sigma^2_{\mathbf{w}^*\hat{\mathbf{w}}}}{\sigma_{\hat{\mathbf{w}}}}}\mathbf{x}_2\right) \\ &= \frac{1}{\pi}\arccos\left(\frac{\sigma_{\mathbf{w}^*\hat{\mathbf{w}}}}{\sqrt{\rho_{\mathbf{w}^\star}\sigma_{\hat{\mathbf{w}}}}}\right)\end{aligned} \tag{7}$$

其中，

$$\rho_{\mathbf{w}^\star} = \sigma_{\mathbf{w}^*} \equiv \lim_{n\to\infty}\mathbb{E}_{\mathbf{w}^\star}\frac{1}{n}\|\mathbf{w}^\star\|_2^2 \tag{8}$$

并且根据 Nishimori 条件，在贝叶斯最优中有

$$\sigma_{\mathbf{w}^\star\hat{\mathbf{w}}} = \lim_{n\to\infty}\mathbb{E}_{\mathbf{w}^\star,\mathrm{X}}\frac{1}{n}\hat{\mathbf{w}}^\top\mathbf{w}^\star = q = m = \sigma_{\hat{\mathbf{w}}} = \lim_{n\to\infty}\mathbb{E}_{\mathbf{w}^\star,\mathrm{X}}\frac{1}{n}\|\hat{\mathbf{w}}\|_2^2 \tag{9}$$

因此泛化误差最终的形式为

$$\varepsilon_{\text{gen}} = \frac{1}{\pi}\arccos\left(\sqrt{\frac{q}{\rho_{\mathbf{w}^\star}}}\right) \tag{10}$$

1.2 主要结果

在 Sec.2 中，我们从 BP 方程出发，推导了 GAMP 算法，总结如下

算法 1: 广义消息传递方程（GAMP）

输入: $\mathbf{y}$ 、$\mathbf{X}$

初始化 $\hat{w}^0$ 、v^0 、f_{out}^0

设置迭代次数 $t=1$，最大迭代次数 T

while $t<T$ **do**

计算 ω 、V

$$\omega_\mu^{t+1} \leftarrow \frac{1}{\sqrt{n}} \sum_i X_{\mu i} \hat{w}_i^t - V_\mu^t f_{\text{out}}^t \left(\omega_\mu^t, y_\mu, V_\mu^t\right)$$

$$V_\mu^{t+1} \leftarrow \frac{1}{n} \sum_i X_{\mu i}^2 v_i^t$$

计算 f_{out}

$$f_{\text{out}}^{t+1} \leftarrow f_{\text{out}}(y, \omega^{t+1}, V^{t+1})$$

其中

$$\mathcal{Z}_{\text{out}^\star}(y,\omega,V) = \mathcal{N}_y\left(1,\Delta^\star\right)\frac{1}{2}\left(1+\operatorname{erf}\left(\frac{\omega}{\sqrt{2V}}\right)\right)+\mathcal{N}_y\left(-1,\Delta^\star\right)\frac{1}{2}\left(1-\operatorname{erf}\left(\frac{\omega}{\sqrt{2V}}\right)\right)$$

计算 Σ 、R

$$\Sigma_i^{t+1} \leftarrow \left[-\frac{1}{n}\sum_\mu X_{\mu i}^2 \partial_\omega f_{\text{out}}\left(\omega_\mu^{t+1}, y_\mu, V_\mu^{t+1}\right)\right]^{-1}$$

$$R_i^{t+1} \leftarrow \hat{w}_i^t + \frac{1}{\sqrt{n}} \left(\Sigma_i\right)^{t+1} \sum_\mu X_{\mu i} f_{\text{out}}\left(\omega_\mu^{t+1}, y_\mu, V_\mu^{t+1}\right)$$

其中

$$\partial_\omega f_{\text{out}}(\omega,y,V) = \frac{\int \mathrm{d}z P_{\text{out}}\left(y|z\right)(z-\omega)^2 e^{-\frac{(z-\omega)^2}{2V}}}{V^2 \int \mathrm{d}z P_{\text{out}}\left(y|z\right) e^{-\frac{(z-\omega)^2}{2V}}} - \frac{1}{V} - f_{\text{out}}^2(\omega,y,V)$$

计算 $\hat{w}$ 、v

$$\hat{w}_i^{t+1} \leftarrow \frac{\Sigma_i^{t+1}}{1+R_i^{t+1}}$$

$$v_i^{t+1} \leftarrow \frac{1}{1+R_i^{t+1}}$$

if $\hat{w}$ 、v 不再变化 **then**

停止迭代

end

$t \leftarrow t+1$

end

输出: $\hat{w}$ 、v

在 Sec.3 中，我们从 GAMP 方程出发推导了状态演化方程（SE），得到

$$
\begin{aligned}
q &= \frac{\hat{q}}{1+\hat{q}} \\
\hat{q} &= \frac{2}{\pi}\frac{\alpha}{1-q}\int D\xi \frac{\exp\left\{-\frac{q\xi^2}{1-q}\right\}}{1+\operatorname{erf}\left(\frac{\sqrt{q}\xi}{\sqrt{2(1-q)}}\right)}
\end{aligned} \tag{11}
$$

在 Sec.4 中，我们使用复本方法进行计算，得到的结果与 SE 一致。我们分别使用 GAMP 和 SE 绘制了泛化误差随数据量密度 α 的曲线，见图 1。在 $\alpha \to \infty$ 时，泛化误差具有以下渐近行为

$$
\varepsilon_{\text{gen}} \simeq \frac{0.4477}{\alpha} \tag{12}
$$

与 SE 方程得到的理论曲线进行比较，见图 2，可以看到当 $\alpha \to \infty$ 时两条曲线基本重合。

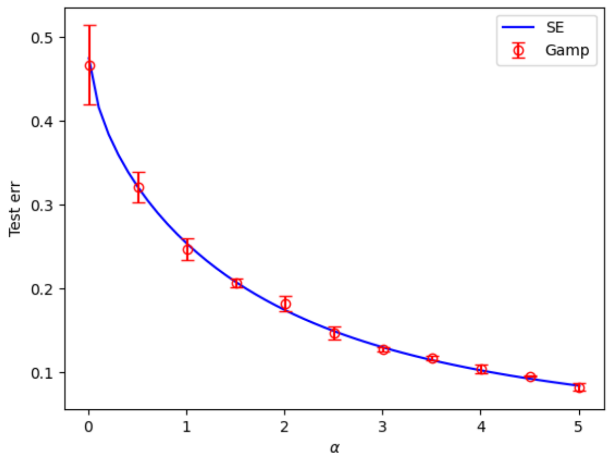

图 1: 泛化误差随数据量密度的变化

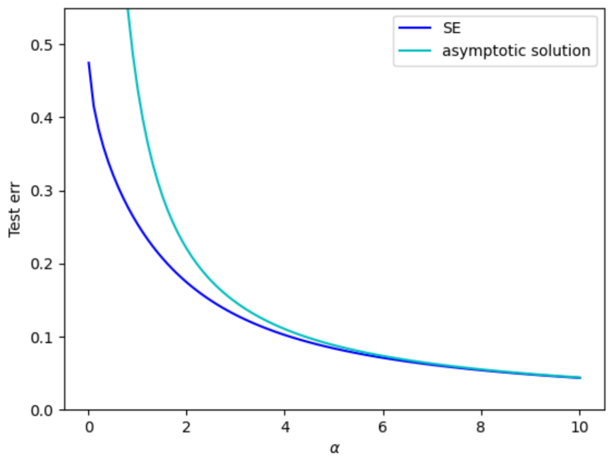

图 2: 理论曲线与渐近解的比较

2 Generalized approximate message passing

根据式 (2) 的因子连乘形式，容易写出 BP 方程：

$$m_{i\to\mu}(w_i) = \frac{1}{z_{i\to\mu}} P_0(w_i) \prod_{\gamma\neq\mu} m_{\gamma\to i}(w_i) \tag{13}$$

$$m_{\mu\to i}(w_i) = \frac{1}{z_{\mu\to i}} \int \prod_{j\neq i} \mathrm{d}w_j P_{\text{out}} \left(y_\mu \mid \frac{1}{\sqrt{n}} \sum_{l=1}^{n} w_l X_{l\mu} \right) m_{j\to\mu}(w_j) \tag{14}$$

为了方便推导，我们在下面的过程中将 $\frac{1}{\sqrt{n}}\mathbf{w}\mathbf{X}$ 中的 $\frac{1}{\sqrt{n}}$ 吸收到 $\mathbf{X}$ 中，最后再根据 $\mathbf{X}$ 的 order 确定各方程的 scaling。

为了导出 relaxed BP 方程，我们假设消息是高斯分布的，$m_{i\to\mu}$ 的均值和方差为

$$\hat{w}_{i\to\mu} \equiv \int \mathrm{d}w_i\, m_{i\to\mu}(w_i)\, w_i \tag{15}$$

$$v_{i\to\mu} \equiv \int \mathrm{d}w_i\, m_{i\to\mu}(x_i)\, w_i^2 - \hat{w}_{i\to\mu}^2 \tag{16}$$

在中心极限定理下有

$$\sum_{j\neq i} X_{\mu j} w_j \sim \mathcal{N}(\omega_{\mu\to i}, V_{\mu\to i}) \tag{17}$$

其中均值和方差分别为

$$\omega_{\mu\to i} = \sum_{j\neq i} X_{\mu j} \hat{w}_{j\to\mu} \tag{18}$$

$$V_{\mu\to i} = \sum_{j\neq i} X_{\mu j}^2 v_{j\to\mu} \tag{19}$$

因此式 (14) 可以改写为

$$m_{\mu\to i}(w_i) \propto \int dz_\mu P_{\text{out}}(y_\mu|z_\mu) \exp\left\{ -\frac{(z-\omega_{\mu\to i}-X_{\mu i}w_i)^2}{2V_{\mu\to i}} \right\} \tag{20}$$

根据展开式

$$(z-\omega_{\mu\to i}-X_{\mu i}w_i)^2 = (z-\omega_{\mu\to i})^2 + X_{\mu i}^2 w_i^2 - 2(z-\omega_{\mu\to i}) X_{\mu i} w_i \tag{21}$$

考虑到 $X_{\mu i} \sim \mathcal{O}(1/\sqrt{n})$，做泰勒展开

$$\begin{aligned}
&\exp\left\{ -\frac{(z-\omega_{\mu\to i}-X_{\mu i}w_i)^2}{2V_{\mu\to i}} \right\} \\
&= \exp\left\{ -\frac{(z-\omega_{\mu\to i})^2 + X_{\mu i}^2 w_i^2 - 2(z-\omega_{\mu\to i}) X_{\mu i} w_i}{2V_{\mu\to i}} \right\} \\
&= \exp\left\{ -\frac{(z-\omega_{\mu\to i})^2}{2V_{\mu\to i}} \right\} \exp\left\{ -\frac{X_{\mu i}^2 w_i^2 - 2(z-\omega_{\mu\to i}) X_{\mu i} w_i}{2V_{\mu\to i}} \right\} \\
&= \exp\left\{ -\frac{(z-\omega_{\mu\to i})^2}{2V_{\mu\to i}} \right\} \left(1 + X_{\mu i}^2 w_i^2 - 2(z-\omega_{\mu\to i}) X_{\mu i} w_i + \frac{1}{2}(z-\omega_{\mu\to i})^2 X_{\mu i}^2 w_i^2 + \mathcal{O}(\frac{1}{n}) \right)
\end{aligned} \tag{22}$$

因此

$$
\begin{aligned}
m_{\mu\to i}(w_i) =& \frac{1}{z}\int dz_\mu P_{\text{out}}(y_\mu|z_\mu)\exp\left\{-\frac{(z-\omega_{\mu\to i})^2}{2V_{\mu\to i}}\right\} \\
& \times\left(1+X_{\mu i}^2 w_i^2-2(z-\omega_{\mu\to i})X_{\mu i}w_i+\frac{1}{2}(z-\omega_{\mu\to i})^2X_{\mu i}^2w_i^2+\mathcal{O}(\frac{1}{d})\right) \\
=& \frac{1}{z}\int dz_\mu P_{\text{out}}(y_\mu|z_\mu)\ \exp\left\{-\frac{(z-\omega_{\mu\to i})^2}{2V_{\mu\to i}}\right\} \\
& +\frac{1}{z}\int dz_\mu P_{\text{out}}(y_\mu|z_\mu)\ X_{\mu i}^2 w_i^2\ \exp\left\{-\frac{(z-\omega_{\mu\to i})^2}{2V_{\mu\to i}}\right\} \\
& -\frac{1}{z}\int dz_\mu P_{\text{out}}(y_\mu|z_\mu)\ 2(z-\omega_{\mu\to i})X_{\mu i}w_i\ \exp\left\{-\frac{(z-\omega_{\mu\to i})^2}{2V_{\mu\to i}}\right\} \\
& +\frac{1}{z}\int dz_\mu P_{\text{out}}(y_\mu|z_\mu)\ \frac{1}{2}(z-\omega_{\mu\to i})^2X_{\mu i}^2w_i^2\ \exp\left\{-\frac{(z-\omega_{\mu\to i})^2}{2V_{\mu\to i}}\right\}
\end{aligned} \tag{23}
$$

定义输出函数

$$
f_{\text{out}}(\omega,y,V)\equiv\frac{\int \mathrm{d}zP_{\text{out}}(y|z)(z-\omega)e^{-\frac{(z-\omega)^2}{2V}}}{V\int \mathrm{d}zP_{\text{out}}(y|z)e^{-\frac{(z-\omega)^2}{2V}}} \tag{24}
$$

并且根据

$$
\frac{\int \mathrm{d}zP_{\text{out}}(y|z)(z-\omega)^2e^{-\frac{(z-\omega)^2}{2V}}}{V^2\int \mathrm{d}zP_{\text{out}}(y|z)e^{-\frac{(z-\omega)^2}{2V}}}=\frac{1}{V}+\partial_\omega f_{\text{out}}(\omega,y,V)+f_{\text{out}}^2(\omega,y,V) \tag{25}
$$

有

$$
\partial_\omega f_{\text{out}}(\omega,y,V)=\frac{\int \mathrm{d}zP_{\text{out}}(y|z)(z-\omega)^2e^{-\frac{(z-\omega)^2}{2V}}}{V^2\int \mathrm{d}zP_{\text{out}}(y|z)e^{-\frac{(z-\omega)^2}{2V}}}-\frac{1}{V}-f_{\text{out}}^2(\omega,y,V) \tag{26}
$$

令

$$
B_{\mu\to i}^t = X_{\mu i}\ f_{\text{out}}\left(\omega_{\mu\to i}^t,y_\mu,V_{\mu\to i}^t\right) \tag{27}
$$

$$
A_{\mu\to i}^t = -X_{\mu i}^2\ \partial_\omega f_{\text{out}}\left(\omega_{\mu\to i}^t,y_\mu,V_{\mu\to i}^t\right) \tag{28}
$$

式 (23) 改写为

$$
m_{\mu\to i}(t,x_i)=\sqrt{\frac{A_{\mu\to i}^t}{2\pi N}}\exp\left\{-\frac{x_i^2}{2N}A_{\mu\to i}^t+B_{\mu\to i}^t\frac{x_i}{\sqrt{N}}-\frac{\left(B_{\mu\to i}^t\right)^2}{2A_{\mu\to i}^t}\right\} \tag{29}
$$

类似的，定义

$$
\Sigma_{\mu\to i}^{t+1}=\frac{1}{\sum_\mu A_{\mu\to i}^{t+1}} \tag{30}
$$

$$
R_{\mu\to i}^{t+1}=\frac{\sum_\mu B_{\mu\to i}^{t+1}}{\sum_\mu A_{\mu\to i}^{t+1}} \tag{31}
$$

式 (13) 改写为

$$
m_{i\to\mu}(w_i)\propto P_0(w_i)\,e^{-\frac{\left(w_i-R_{i\to\mu}\right)^2}{2\Sigma_{i\to\mu}}} \tag{32}
$$

定义函数

$$
f_{\text{w}}\equiv\frac{\int \mathrm{d}w\ wP_0(w)e^{-\frac{(w-R)^2}{2\Sigma}}}{\int \mathrm{d}w\ P_0(w)e^{-\frac{(w-R)^2}{2\Sigma}}}=\frac{\Sigma}{1+R} \tag{33}
$$

$\hat{w}$ 和 v 通过下式更新

$$\hat{w}_{\mu\to i} = f_{\mathrm{w}}(\Sigma, R) = \frac{\Sigma}{1+R} \tag{34}$$

$$v_{\mu\to i} = \partial_R f_{\mathrm{w}}(\Sigma, R) = \frac{1}{1+R} \tag{35}$$

至此我们推导出了 r-BP 方程，下面继续化简推导 GAMP 方程。

注意到式 (19) 中 $X^2 \sim \mathcal{O}(1/n)$，说明 V 的连接比较弱，做近似

$$V_\mu^{t+1} = \sum_i X_{\mu i}^2 v_{i\to\mu}^t \approx \sum_i X_{\mu i}^2 v_i^t \tag{36}$$

式 (30) 可以化简为

$$(\Sigma_i)^{t+1} = \left[-\sum_\mu X_{\mu i}^2 \partial_\omega f_{\mathrm{out}}\left(\omega_\mu^{t+1}, y_\mu, V_\mu^{t+1}\right)\right]^{-1} \tag{37}$$

输出函数 (24) 化简为

$$\begin{aligned} f_{\mathrm{out}}\left(\omega_{\mu\to i}^{t+1}, y_\mu, V_{\mu\to i}^{t+1}\right) &\approx f_{\mathrm{out}}\left(\omega_\mu^{t+1}, y_\mu, V_\mu^{t+1}\right) - X_{\mu i}\hat{w}_{i\to\mu}^t \partial_\omega f_{\mathrm{out}}\left(\omega_\mu^{t+1}, y_\mu, V_\mu^{t+1}\right) \\ &\approx f_{\mathrm{out}}\left(\omega_\mu^{t+1}, y_\mu, V_\mu^{t+1}\right) - X_{\mu i}\hat{w}_i^t \partial_\omega f_{\mathrm{out}}\left(\omega_\mu^{t+1}, y_\mu, V_\mu^{t+1}\right) \end{aligned} \tag{38}$$

因此式 (31) 可以化简为

$$\begin{aligned} R_i^{t+1} &= (\Sigma_i)^{t+1} \times \left[\sum_\mu X_{\mu i} f_{\mathrm{out}}\left(\omega_\mu^{t+1}, y_\mu, V_\mu^{t+1}\right) - X_{\mu i}^2 \hat{w}_i^t \partial_\omega f_{\mathrm{out}}\left(\omega_\mu^{t+1}, y_\mu, V_\mu^{t+1}\right)\right] \\ &= \hat{w}_i^t + (\Sigma_i)^{t+1} \sum_\mu X_{\mu i} f_{\mathrm{out}}\left(\omega_\mu^{t+1}, y_\mu, V_\mu^{t+1}\right) \end{aligned} \tag{39}$$

式 (34) 可以化简为

$$\begin{aligned} \hat{w}_{i\to\mu}^t &= f_{\mathrm{w}}\left(R_{i\to\mu}^t, \Sigma_{i\to\mu}\right) \approx f_{\mathrm{w}}\left(R_{i\to\mu}^t, \Sigma_i\right) \\ &\approx f_{\mathrm{w}}\left(R_i^t, \Sigma_i\right) - B_{\mu\to i}^t \partial_R f_{\mathrm{w}}\left(R_i^t, \Sigma_i\right) \\ &\approx \hat{w}_i^t - f_{\mathrm{out}}\left(\omega_\mu^t, y_\mu, V_\mu^t\right) X_{\mu i} v_i^t \end{aligned} \tag{40}$$

因此式 (18) 化简为

$$\omega_\mu^{t+1} = \sum_i X_{\mu i}\hat{w}_i^t - \sum_i f_{\mathrm{out}}\left(\omega_\mu^t, y_\mu, V_\mu^t\right) X_{\mu i}^2 v_i^t = \sum_i X_{\mu i}\hat{w}_i - V_\mu^t f_{\mathrm{out}}\left(\omega_\mu^t, y_\mu, V_\mu^t\right) \tag{41}$$

GAMP 推导完毕，我们将 $\mathbf{X}$ 吸收的 scaling $1/\sqrt{n}$ 重新标注出来，并将各式总结如下

初始化 $\hat{w}_i^0, v_i^0, f_{\text{out}}^0$

$$
\begin{aligned}
V_\mu^{t+1} &= \frac{1}{n}\sum_i X_{\mu i}^2 v_i^t \\
\omega_\mu^{t+1} &= \frac{1}{\sqrt{n}}\sum_i X_{\mu i}\hat{w}_i^t - V_\mu^t f_{\text{out}}^t \\
f_{\text{out}}^{t+1} &= f_{\text{out}}(y, \omega^{t+1}, V^{t+1}) \\
\Sigma_i^{t+1} &= \left[-\frac{1}{n}\sum_\mu X_{\mu i}^2 \partial_\omega f_{\text{out}}\left(\omega_\mu^{t+1}, y_\mu, V_\mu^{t+1}\right)\right]^{-1} \\
R_i^{t+1} &= \hat{w}_i^t + \frac{1}{\sqrt{n}}\left(\Sigma_i\right)^{t+1}\sum_\mu X_{\mu i} f_{\text{out}}\left(\omega_\mu^{t+1}, y_\mu, V_\mu^{t+1}\right) \\
\hat{w}_i^{t+1} &= \frac{\Sigma_i^{t+1}}{1+R_i^{t+1}} \\
v_i^{t+1} &= \frac{1}{1+R_i^{t+1}}
\end{aligned}
$$

其中，

$$
\begin{aligned}
\mathcal{Z}_{\text{out}^\star}(y,\omega,V) &= \mathcal{N}_y\left(1,\Delta^\star\right)\frac{1}{2}\left(1+\operatorname{erf}\left(\frac{\omega}{\sqrt{2V}}\right)\right) + \mathcal{N}_y\left(-1,\Delta^\star\right)\frac{1}{2}\left(1-\operatorname{erf}\left(\frac{\omega}{\sqrt{2V}}\right)\right) \\
\partial_\omega f_{\text{out}}(\omega,y,V) &= \frac{\int \mathrm{d}z P_{\text{out}}\left(y|z\right)(z-\omega)^2 e^{-\frac{(z-\omega)^2}{2V}}}{V^2\int \mathrm{d}z P_{\text{out}}\left(y|z\right) e^{-\frac{(z-\omega)^2}{2V}}} - \frac{1}{V} - f_{\text{out}}^2(\omega,y,V)
\end{aligned}
$$

3 State evolution equation

在本小节中，我们展示了如何从 GAMP 方程推导 state evolution 方程，主要参考了[1]
首先, 我们明确我们的 channel 输出概率分布为

$$P_{out}(y|z) = \delta[y - \mathrm{sign}(z)] \tag{42}$$

并且我们回顾 GAMP 算法中的这两个让人又爱又恨的函数:

$$f_{\mathrm{w}}(\Sigma, R) \equiv \frac{\int \mathrm{d}w\ wP_w(w)e^{-\frac{(w-R)^2}{2\Sigma}}}{\int \mathrm{d}w\ P_w(w)e^{-\frac{(w-R)^2}{2\Sigma}}} \tag{43}$$

$$f_{\mathrm{out}}(\omega, y, V) \equiv \frac{\int \mathrm{d}zP_{\mathrm{out}}(y|z)(z-\omega)e^{-\frac{(z-\omega)^2}{2V}}}{V\int \mathrm{d}zP_{\mathrm{out}}(y|z)e^{-\frac{(z-\omega)^2}{2V}}} \tag{44}$$

3.1 模型出发点:

其次，我们关注 GAMP 算法中引入的和模型有关的量

$$\omega_{\mu\to i} = \frac{1}{\sqrt{n}}\sum_{j\neq i} X_{\mu i}\hat{w}_i \tag{45}$$

$$z_{\mu\to i} = \frac{1}{\sqrt{n}}\sum_{j\neq i} X_{\mu i}w^*_{i\to j} \tag{46}$$

$$V_\mu = \frac{1}{n}\sum_i X^2_{\mu i}v_i \tag{47}$$

$$\tag{48}$$

在 N 趋于无穷时，得到:

$$V_\mu = \frac{1}{n}\sum_i v_i \tag{49}$$

因此我们定义以下序参量:

$$q = \mathbb{E}[\omega^2] = \mathbb{E}[\hat{w}^2] \tag{50}$$

$$m = \mathbb{E}[z\omega] = \mathbb{E}[w^*\hat{w}] \tag{51}$$

$$\tag{52}$$

3.2 算法出发点:

接下来，我们关注 GAMP 算法中引入的和算法相关的量

$$\Sigma_i = \frac{1}{\sum_\mu A_{\mu\to i}} \tag{53}$$

$$R_i = \frac{\sum_\mu B_{\mu\to i}}{\sum_\mu A_{\mu\to i}} \tag{54}$$

我们考虑计算：

$$\begin{aligned}\frac{R_i}{\Sigma_i} &= \sum_\mu B_{\mu\to i} \\ &= \sum_\mu X_{\mu i} f_{\text{out}}(\omega_{\mu\to i}, y_\mu, V_{\mu\to i}) \\ &= \sum_\mu X_{\mu i} f_{\text{out}}(\omega_{\mu\to i}, \text{sign}[\sum_{j\neq i} X_{\mu j} w_j^* + X_{\mu i} w_i^*], V) \\ &= \sum_\mu X_{\mu i} f_{\text{out}}(\omega_{\mu\to i}, \text{sign}[\sum_{j\neq i} X_{\mu j} w_j^*], V) + \sum_\mu X_{\mu i} f_{\text{out}}(\omega_{\mu\to i}, \text{sign}[X_{\mu i} w_i^*], V)\end{aligned} \tag{55}$$

在这里我们继续定义了新的参数[1].：

$$\hat{q} = \alpha \mathbb{E}_{\omega,z}[f_{\text{out}}^2(\omega, \text{sign}[z], V)] \tag{56}$$

$$\hat{m} = \alpha \mathbb{E}_{\omega,z}[\partial_z f_{\text{out}}(\omega, \text{sign}[z], V] \tag{57}$$

故，上式可以写成：

$$\frac{R_i}{\Sigma_i} = \mathcal{N}(0,1)\sqrt{\hat{q}} + w_i^* \hat{m} \tag{58}$$

3.3 态演化方程显式表达：

根据序参量的定义，我们仔细写出平均, 从而闭合方程：

$$q = \mathbb{E}_{w^*} \mathbb{E}_{R,\Sigma}[f_{\text{w}}^2(\Sigma, R)] \tag{59}$$

$$m = \mathbb{E}_{w^*} \mathbb{E}_{R,\Sigma}[w^* f_{\text{w}}(\Sigma, R)] \tag{60}$$

最后在贝叶斯最优的框架下，再有 Nishimori 条件 $q = m$，我们得到显示态演化方程显式表达：

$$q^{t+1} = \int dx P_X(x) \int d\xi \frac{e^{-\frac{\xi^2}{2}}}{\sqrt{2\pi}} f_{w^*}^2(\frac{1}{\hat{q}^t}, x + \frac{\xi}{\sqrt{\hat{q}^t}}) \tag{61}$$

$$\hat{q}^t = -\int dp \int dz \frac{e^{-\frac{p^2}{2m^t}} e^{-\frac{(z-p)^2}{2(1-m^t)}}}{2\pi\sqrt{m^t(1-m^t)}} \partial_p f_{\text{out}}(p, \text{sign}[z], 1-m^t) \tag{62}$$

最后通过一系列繁琐的高斯积分计算, 利用下面四个公式纵横捭阖

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$$

$$\int_{-\infty}^{\infty} x e^{-x^2} dx = 0$$

$$\int_{-\infty}^{\infty} x^2 e^{-x^2} dx = \frac{\sqrt{\pi}}{2}$$

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$$

---

[1]我们这里定义的 $\hat{m}\hat{q}$ 与[1]不太一样，
是为了可以和其他论文推导如[2]得到态演化方程一致

最终得到：

$$
\begin{aligned}
q &= \frac{\hat{q}}{1+\hat{q}} \\
\hat{q} &= \frac{2}{\pi}\frac{\alpha}{1-q}\int D\xi \frac{\exp\left\{-\frac{q\xi^2}{1-q}\right\}}{1+\operatorname{erf}\left(\frac{\sqrt{q}\xi}{\sqrt{2(1-q)}}\right)}
\end{aligned}
\tag{63}
$$

4 Replica symmetry analysis

在贝叶斯最优框架（式 2）下，我们可以写出配分函数

$$\mathcal{Z}(\mathbf{y},\mathbf{X}) = P(\mathbf{y},\mathbf{X}) = \int \mathrm{d}\mathbf{w}\ P(\mathbf{w},\mathbf{y},\mathbf{X}) = \int \mathrm{d}\mathbf{w}\ P(\mathbf{y}|\mathbf{w},\mathbf{X})P(\mathbf{w}) \tag{64}$$

为了方便，引入变量 $\mathbf{z} = \mathbf{w}\mathbf{X}/\sqrt{n}$，因此有

$$P(\mathbf{y}|\mathbf{z}) = \int \mathrm{d}\mathbf{z}\ P(\mathbf{y}|\mathbf{w},\mathbf{X})\,\delta\left(\mathbf{z} - \frac{1}{\sqrt{n}}\mathbf{w}\mathbf{X}\right) \tag{65}$$

式 (64) 改写为

$$\mathcal{Z}(\mathbf{y},\mathbf{X}) = \int \mathrm{d}\mathbf{z}\ P(\mathbf{y}|\mathbf{z}) \int \mathrm{d}\mathbf{w}\ P(\mathbf{w})\,\delta\left(\mathbf{z} - \frac{1}{\sqrt{n}}\mathbf{w}\mathbf{X}\right) \tag{66}$$

为了求解自由能的淬火平均

$$\Phi = \frac{1}{n}\mathbb{E}_{\mathbf{y},\mathbf{X}}\ \log \mathcal{Z}(\mathbf{y},\boldsymbol{X}) \tag{67}$$

使用复本方法将对 $\log \mathcal{Z}(\mathbf{y},\boldsymbol{X})$ 的平均转为对 $\mathcal{Z}(\mathbf{y},\boldsymbol{X})^r$ 的平均

$$\Phi = \frac{1}{n}\lim_{r\to 0}\frac{\partial \log \mathbb{E}_{\mathbf{y},\mathbf{X}}\ \left[\mathcal{Z}(\mathbf{y},\boldsymbol{X})^r\right]}{\partial r} \tag{68}$$

其中 r 是复本指标。下面求解 $\mathbb{E}_{\mathbf{y},\mathbf{X}}\ \left[\mathcal{Z}(\mathbf{y},\boldsymbol{X})^r\right]$

$$\begin{aligned}
&\mathbb{E}_{\mathbf{y},\mathbf{X}}\left[\mathcal{Z}(\mathbf{y},\mathbf{X})^r\right] \\
&= \mathbb{E}_{\mathbf{w}^\star,\mathbf{X}}\left[\prod_{a=1}^{r}\int_{\mathbb{R}^n}\mathrm{d}\mathbf{z}^a P_{\mathrm{out}^a}\left(\mathbf{y}\mid\mathbf{z}^a\right)\int_{\mathbb{R}^d}\mathrm{d}\mathbf{w}^a P_{\mathbf{w}^a}\left(\mathbf{w}^a\right)\delta\left(\mathbf{z}^a - \frac{1}{\sqrt{n}}\mathbf{X}\mathbf{w}^a\right)\right] \\
&= \mathbb{E}_{\mathbf{X}}\int_{\mathbb{R}^n}\mathrm{d}\mathbf{y}\int_{\mathbb{R}^n}\mathrm{d}\mathbf{z}^\star P_{\mathrm{out}^\star}\left(\mathbf{y}\mid\mathbf{z}^\star\right)\int_{\mathbb{R}^d}\mathrm{d}\mathbf{w}^\star P_{\mathbf{w}^\star}\left(\mathbf{w}^\star\right)\delta\left(\mathbf{z}^\star - \frac{1}{\sqrt{n}}\mathbf{X}\mathbf{w}^\star\right) \\
&\times\left[\prod_{a=1}^{r}\int_{\mathbb{R}^n}\mathrm{d}\mathbf{z}^a P_{\mathrm{out}^a}\left(\mathbf{y}\mid\mathbf{z}^a\right)\int_{\mathbb{R}^d}\mathrm{d}\mathbf{w}^a P_{\mathbf{w}^a}\left(\mathbf{w}^a\right)\delta\left(\mathbf{z}^a - \frac{1}{\sqrt{n}}\mathbf{X}\mathbf{w}^a\right)\right] \\
&= \mathbb{E}_{\mathbf{X}}\int_{\mathbb{R}^n}\mathrm{d}\mathbf{y}\prod_{a=0}^{r}\int_{\mathbb{R}^n}\mathrm{d}\mathbf{z}^a P_{\mathrm{out}^a}\left(\mathbf{y}\mid\mathbf{z}^a\right)\int_{\mathbb{R}^d}\mathrm{d}\mathbf{w}^a P_{\mathbf{w}^a}\left(\mathbf{w}^a\right)\delta\left(\mathbf{z}^a - \frac{1}{\sqrt{n}}\mathbf{X}\mathbf{w}^a\right)
\end{aligned} \tag{69}$$

假设所有的数据 X 都是独立同分布的，根据中心极限定理，有

$$z_\mu^a = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} x_i^{(\mu)} w_i^a \sim \mathcal{N}\left(\mathbb{E}_{\mathrm{X}}\left[z_\mu^a\right], \mathbb{E}_{\mathrm{X}}\left[z_\mu^a z_\mu^b\right]\right) \tag{70}$$

其中

$$\begin{aligned}
\mathbb{E}_{\mathrm{X}}\left[z_\mu^a\right] &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbb{E}_{\mathrm{X}}\left[x_i^{(\mu)}\right] w_i^a = 0 \\
\mathbb{E}_{\mathrm{X}}\left[z_\mu^a z_\mu^b\right] &= \frac{1}{n}\sum_{ij}\mathbb{E}_{\mathrm{X}}\left[x_i^{(\mu)} x_j^{(\mu)}\right] w_i^a w_j^b = \frac{1}{n}\sum_{ij}\delta_{ij} w_i^a w_j^b = \frac{1}{n}\mathbf{w}^a\cdot\mathbf{w}^b \equiv Q
\end{aligned} \tag{71}$$

定义对称重叠矩阵（symmetric overlap matrix）

$$Q\left(\{\mathbf{w}^a\}\right) \equiv \left(\frac{1}{n}\mathbf{w}^a \cdot \mathbf{w}^b\right)_{a,b=0..r} \tag{72}$$

以及 $\tilde{\mathbf{z}}_\mu \equiv \left(z_\mu^a\right)_{a=0\cdots r}$，$\tilde{\mathbf{w}}_i \equiv \left(w_i^a\right)_{a=0\cdots r}$，因此有

$$\tilde{\mathbf{z}}_\mu \sim P_{\tilde{\mathbf{z}}}(\tilde{\mathbf{z}}; Q) = \mathcal{N}_{\tilde{\mathbf{z}}}\left(\mathbf{0}_{r+1}, Q\right) \qquad P_{\tilde{\mathrm{w}}}(\tilde{\mathbf{w}}) = \prod_{a=0}^{r} P_{\mathrm{w}}\left(\tilde{w}^a\right) \tag{73}$$

式 (69) 可以继续化简为

$$\begin{aligned}\mathbb{E}_{\mathbf{y},\mathbf{X}}\left[\mathcal{Z}(\mathbf{y},\mathbf{X})^r\right] &= \mathbb{E}_{\mathbf{X}} \int \mathrm{d}\mathbf{y} \prod_{a=0}^{r} \int \mathrm{d}\mathbf{z}^a P_{\mathrm{out}^a}\left(\mathbf{y} \mid \mathbf{z}^a\right) \int \mathrm{d}\mathbf{w}^a P_{\mathbf{w}^a}\left(\mathbf{w}^a\right) \delta\left(\mathbf{z}^a - \frac{1}{\sqrt{d}}\mathbf{X}\mathbf{w}^a\right) \\ &= \left[\int \mathrm{d}y \int \mathrm{d}\tilde{\mathbf{z}} P_{\mathrm{out}}(y \mid \tilde{\mathbf{z}}) P_{\tilde{\mathbf{z}}}(\tilde{\mathbf{z}}; Q(\tilde{\mathbf{w}}))\right]^d \left[\int \mathrm{d}\tilde{\mathbf{w}} P_{\tilde{\mathbf{w}}}(\tilde{\mathbf{w}})\right]^n\end{aligned} \tag{74}$$

利用 $\delta(x)$ 的傅里叶变换

$$\begin{aligned} 1 &= \int \mathrm{d}Q \prod_{a\le b} \delta\left(nQ_{ab} - \sum_{i=1}^{n} w_i^a w_i^b\right) \\ &\propto \int \mathrm{d}Q \int \mathrm{d}\hat{Q} \prod_{a\le b} \exp\left\{-\hat{Q}_{ab}\left(nQ_{ab} - \sum_{i=1}^{n} w_i^a w_i^b\right)\right\} \\ &\propto \int \mathrm{d}Q \int \mathrm{d}\hat{Q} \exp\left\{-\sum_{a\le b} \hat{Q}_{ab}\left(nQ_{ab} - \sum_{i=1}^{n} w_i^a w_i^b\right)\right\} \\ &\propto \int \mathrm{d}Q \int \mathrm{d}\hat{Q} \exp(-n\,\mathrm{Tr}[Q\hat{Q}]) \exp\left(\frac{1}{2}\sum_{i=1}^{n} \tilde{\mathbf{w}}_i^\top \hat{Q} \tilde{\mathbf{w}}_i\right) \end{aligned} \tag{75}$$

可以将式 (74) 转为对 Q 和 $\hat{Q}$ 的积分，并使用 Laplace 近似有

$$\begin{aligned}\mathbb{E}_{\mathbf{y},\mathbf{X}}\left[\mathcal{Z}(\mathbf{y},\mathbf{X})^r\right] \propto& \int \mathrm{d}Q \int \mathrm{d}\hat{Q}\ \exp(-n\,\mathrm{Tr}[Q\hat{Q}]) \exp\left(\frac{1}{2}\sum_{i=1}^{n} \tilde{\mathbf{w}}_i^\top \hat{Q} \tilde{\mathbf{w}}_i\right) \\ & \left[\int \mathrm{d}y \int \mathrm{d}\tilde{\mathbf{z}}\ P_{\mathrm{out}}\left(y \mid \tilde{\mathbf{z}}\right) P_{\tilde{\mathbf{z}}}(\tilde{\mathbf{z}}; Q(\tilde{\mathbf{w}}))\right]^d \left[\int \mathrm{d}\tilde{\mathbf{w}} P_{\tilde{\mathbf{w}}}(\tilde{\mathbf{w}})\right]^n \\ \propto& \iint \mathrm{d}Q\ \mathrm{d}\hat{Q}\ e^{n\Phi^{(r)}(Q,\hat{Q})}\end{aligned} \tag{76}$$

其中，

$$\Phi^{(r)}(Q,\hat{Q}) = -\,\mathrm{Tr}[Q\hat{Q}] + \log \Psi_{\mathrm{w}}^{(r)}(\hat{Q}) + \alpha \log \Psi_{\mathrm{out}}^{(r)}(Q) \tag{77}$$

$$\Psi_{\mathrm{w}}^{(r)}(\hat{Q}) = \int \mathrm{d}\tilde{\mathbf{w}}\ P_{\tilde{\mathbf{w}}}(\tilde{\mathbf{w}}) e^{\frac{1}{2}\tilde{\mathbf{w}}\hat{Q}\tilde{\mathbf{w}}} \tag{78}$$

$$\Psi_{\mathrm{out}}^{(r)}(Q) = \int \mathrm{d}y \int \mathrm{d}\tilde{\mathbf{z}}\ P_{\tilde{z}}(\tilde{\mathbf{z}}; Q) P_{\mathrm{out}}\left(y \mid \tilde{\mathbf{z}}\right) \tag{79}$$

以及，

$$P_{\tilde{z}}(\tilde{\mathbf{z}};Q)=\frac{e^{-\frac{1}{2}\tilde{\mathbf{z}}^{\mathrm{T}}Q^{-1}\tilde{\mathbf{z}}}}{\det(2\pi Q)^{1/2}} \tag{80}$$

取 $r\to 0$ 和 $n\to\infty$ 得到鞍点方程

$$\Phi(\alpha)=\mathrm{extr}_{Q,\hat{Q}}\left\{\lim_{r\to 0}\frac{\partial\Phi^{(r)}(Q,\hat{Q})}{\partial r}\right\} \tag{81}$$

使用复本对称假设，有

$$Q_{\mathrm{rs}}=\begin{pmatrix} Q^0 & m & \dots & m\\ m & Q & \dots & \dots\\ \dots & \dots & \dots & q\\ m & \dots & q & Q\end{pmatrix}\quad \text{and}\quad \hat{Q}_{\mathrm{rs}}=\begin{pmatrix} \hat{Q}^0 & \hat{m} & \dots & \hat{m}\\ \hat{m} & -\frac{1}{2}\hat{Q} & \dots & \dots\\ \dots & \dots & \dots & \hat{q}\\ \hat{m} & \dots & \hat{q} & -\frac{1}{2}\hat{Q}\end{pmatrix} \tag{82}$$

其中

$$\begin{aligned} m&=\frac{1}{n}\mathbf{w}^a\cdot\mathbf{w}^\star\\ q&=\frac{1}{n}\mathbf{w}^a\cdot\mathbf{w}^b\\ Q&=\frac{1}{n}||\mathbf{w}^a||_2^2\\ Q^0&=\rho_{\mathbf{w}^\star}=\frac{1}{n}||\mathbf{w}^\star||_2^2\end{aligned}$$

其中复本的第零项为老师模型，这是由于式 (69) 中求解 $\mathbb{E}_{\mathbf{y},\mathbf{X}}\ [\mathcal{Z}(\mathbf{y},\boldsymbol{X})^r]$ 时对 $\mathbf{y}$ 的平均实际上是对老师的权重 $\mathbf{w}^\star$ 的平均。分别计算式（77）中的三项：

第一项

$$\mathrm{Tr}(Q\hat{Q})\Big|_{\mathrm{rs}}=Q^0\hat{Q}^0+rm\hat{m}-\frac{1}{2}rQ\hat{Q}+\frac{r(r-1)}{2}q\hat{q} \tag{83}$$

因此

$$\begin{aligned}&\mathrm{extr}\left\{\lim_{r\to 0}\frac{\partial}{\partial r}\left(-\mathrm{Tr}[Q\hat{Q}]\right)\right\}\\ &=\mathrm{extr}\left\{\lim_{r\to 0}\frac{\partial}{\partial r}\left(-Q^0\hat{Q}^0-rm\hat{m}+\frac{1}{2}rQ\hat{Q}-\frac{r(r-1)}{2}q\hat{q}\right)\right\}\\ &=\mathrm{extr}\left\{\lim_{r\to 0}\left(-m\hat{m}+\frac{1}{2}Q\hat{Q}-\left(r-\frac{1}{2}\right)q\hat{q}\right)\right\}\\ &=\mathrm{extr}\left\{-m\hat{m}+\frac{1}{2}Q\hat{Q}+\frac{1}{2}q\hat{q}\right\}\end{aligned} \tag{84}$$

第二项 $\Psi_{\mathrm{w}}^{(r)}(\hat{Q})\Big|_{\mathrm{rs}}=\int\mathbf{d}\tilde{\mathbf{w}}P_{\tilde{\mathbf{w}}}(\tilde{\mathbf{w}})\exp\left\{\frac{1}{2}\tilde{\mathbf{w}}\hat{Q}_{\mathrm{rs}}\tilde{\mathbf{w}}\right\}$，其中

$$\tilde{\mathbf{w}}\hat{Q}_{\mathrm{rs}}\tilde{\mathbf{w}}=w^\star\hat{Q}^0w^\star+2\sum_{a=1}^r w^\star\hat{m}w^a-\left(\hat{Q}+\hat{q}\right)\sum_{a=1}^r(w^a)^2+\hat{q}\left(\sum_{a=1}^r w^a\right)^2$$

因此

$$
\begin{aligned}
\Psi_{\mathrm{w}}^{(r)}(\hat{Q})\Big|_{\mathrm{rs}} &= \int \mathbf{d}\tilde{\mathbf{w}} P_{\tilde{\mathbf{w}}}(\tilde{\mathbf{w}}) e^{\frac{1}{2}\tilde{\mathbf{w}}\hat{Q}_{\mathrm{rs}}\tilde{\mathbf{w}}} \\
&= \mathbb{E}_{w^\star} e^{\frac{1}{2}\hat{Q}^0(w^\star)^2} \int \mathbf{d}\tilde{\mathbf{w}} P_{\tilde{\mathbf{w}}}(\tilde{\mathbf{w}}) e^{w^\star \hat{m} \sum_{a=1}^r w^a - \frac{1}{2}(\hat{Q}+\hat{q})\sum_{a=1}^r (\tilde{w}^a)^2 + \frac{1}{2}\hat{q}\left(\sum_{a=1}^r w^a\right)^2} \\
&= \mathbb{E}_{\xi, w^\star} e^{\frac{1}{2}\hat{Q}^0(w^\star)^2} \left[\mathbb{E}_w \exp\left(\left[\hat{m} w^\star w - \frac{1}{2}(\hat{Q}+\hat{q})w^2 + \hat{q}^{1/2}\xi w\right]\right)\right]^r.
\end{aligned} \tag{85}
$$

因此

$$
\begin{aligned}
&\operatorname{extr}\left\{\lim_{r\to 0}\frac{\partial}{\partial r}\left(\log \Psi_{\mathrm{w}}^{(r)}(\hat{Q})\right)\right\} \\
&= \operatorname{extr}\left\{\lim_{r\to 0}\frac{\partial}{\partial r}\mathbb{E}_{\xi, w^\star} r \log\left[\mathbb{E}_w \exp\left(\hat{m} w^\star w - \frac{1}{2}(\hat{Q}+\hat{q})w^2 + \hat{q}^{1/2}\xi w\right)\right]\right\} \\
&= \operatorname{extr}\left\{\mathbb{E}_{\xi, w^\star} \log\left[\mathbb{E}_w \exp\left(\hat{m} w^\star w - \frac{1}{2}(\hat{Q}+\hat{q})w^2 + \hat{q}^{1/2}\xi w\right)\right]\right\}
\end{aligned} \tag{86}
$$

为了解耦老师和学生的项，做变量替换

$$
\xi \leftarrow \xi + \hat{q}^{-\frac{1}{2}}\hat{m} w^\star
$$

可得

$$
\begin{aligned}
&\operatorname{extr}\left\{\lim_{r\to 0}\frac{\partial}{\partial r}\left(\log \Psi_{\mathrm{w}}^{(r)}(\hat{Q})\right)\right\} \\
&= \operatorname{extr}\left\{\mathbb{E}_{\xi, w^\star} \exp\left(-\frac{1}{2}\hat{q}^{-1}\hat{m}^2(w^\star)^2 + \xi\hat{q}^{-\frac{1}{2}}\hat{m} w^\star\right)\log\left[\mathbb{E}_w \exp\left(-\frac{1}{2}(\hat{Q}+\hat{q})w^2 + \hat{q}^{1/2}\xi w\right)\right]\right\}
\end{aligned} \tag{87}
$$

注意到 $P_{\mathbf{w}} \sim \mathcal{N}(0,1)$，定义

$$
\mathcal{Z}_{\mathrm{w}}(\gamma, \Lambda) \equiv \mathbb{E}_{w\sim P_{\mathrm{w}}}\left[e^{-\frac{1}{2}\Lambda w^2 + \gamma w}\right] \tag{88}
$$

类似地，可以定义 $\mathcal{Z}_{\mathrm{w}^\star}$，因此

$$
\operatorname{extr}\left\{\lim_{r\to 0}\frac{\partial}{\partial r}\left(\log \Psi_{\mathrm{w}}^{(r)}(\hat{Q})\right)\right\} = \mathbb{E}_\xi\left[\mathcal{Z}_{\mathrm{w}^\star}\left(\hat{m}\hat{q}^{-1/2}\xi, \hat{m}\hat{q}^{-1}\hat{m}\right)\log \mathcal{Z}_{\mathrm{w}}\left(\hat{q}^{1/2}\xi, \hat{Q}+\hat{q}\right)\right] \tag{89}
$$

第三项，注意到

$$
Q_{\mathrm{rs}}^{-1} = \begin{bmatrix}
Q_{00}^{-1} & Q_{01}^{-1} & Q_{01}^{-1} & Q_{01}^{-1} \\
Q_{01}^{-1} & Q_{11}^{-1} & Q_{12}^{-1} & Q_{12}^{-1} \\
Q_{01}^{-1} & Q_{12}^{-1} & Q_{11}^{-1} & Q_{12}^{-1} \\
Q_{01}^{-1} & Q_{12}^{-1} & Q_{12}^{-1} & Q_{11}^{-1}
\end{bmatrix}
$$

并且

$$\begin{aligned}
Q_{00}^{-1} &= \left(Q^0 - rm(Q+(r-1)q)^{-1}m\right)^{-1} \\
Q_{01}^{-1} &= -\left(Q^0 - rm(Q+(r-1)q)^{-1}m\right)^{-1} m(q+(r-1)q)^{-1} \\
Q_{11}^{-1} &= (Q-q)^{-1} - (Q+(r-1)q)^{-1}q(Q-q)^{-1} \\
&\quad + (Q+(r-1)q)^{-1}m\left(Q^0 - rm(Q+(r-1)q)^{-1}m\right)^{-1} m(Q+(r-1)q)^{-1} \\
Q_{12}^{-1} &= -(Q+(r-1)q)^{-1}q(Q-q)^{-1} \\
&\quad + (Q+(r-1)q)^{-1}m\left(Q - rm(Q+(r-1)q)^{-1}m\right)^{-1} m(Q+(r-1)q)^{-1}
\end{aligned}$$

可以求得

$$\det Q_{\text{rs}} = (Q-q)^{r-1}(Q+(r-1)q)\left(Q^0 - rm(Q+(r-1)q)^{-1}m\right)$$

因此

$$\begin{aligned}
\Psi_{\text{out}}^{(r)}(Q)\Big|_{\text{rs}} &= \int \mathrm{d}y \int \mathrm{d}\tilde{\mathbf{z}} e^{-\frac{1}{2}\tilde{\mathbf{z}}^\top Q_{\text{rs}}^{-1}\tilde{\mathbf{z}} - \frac{1}{2}\log(\det(2\pi Q_{\text{rs}}))} P_{\text{out}}(y \mid \tilde{\mathbf{z}}) \\
&= \mathbb{E}_{y,\xi} \mathrm{e}^{-\frac{1}{2}\log(\det(2\pi Q_{\text{rs}}))} \\
&\times \int \mathrm{d}z^\star P_{\text{out}^\star}(y \mid z^\star)\, e^{-\frac{1}{2}Q_{00}^{-1}(z^\star)^2} \left[\int dz P_{\text{out}}(y \mid z) e^{-Q_{01}^{-1}z^\star z - \frac{1}{2}\left(Q_{11}^{-1} - Q_{12}^{-1}\right)z^2 - Q_{12}^{-1/2}\xi z}\right]^r
\end{aligned} \tag{90}$$

采取与第二项类似的方式，定义 $\mathcal{Z}_{\text{out}}$，可得

$$\operatorname{extr}\left\{\lim_{r\to 0}\frac{\partial}{\partial r}\left(\log \Psi_{\text{out}}^{(r)}(\hat{Q})\right)\right\} = \mathbb{E}_{y,\xi}\left[\mathcal{Z}_{\text{out}^\star}\left(y, mq^{-1/2}\xi, \rho_{\text{w}^\star} - mq^{-1}m\right)\log \mathcal{Z}_{\text{out}}\left(y, q^{1/2}\xi, Q-q\right)\right] \tag{91}$$

综上,

$$\begin{aligned}
\Phi(\alpha) &= \operatorname{extr}_{Q,\hat{Q}}\left\{\lim_{r\to 0}\frac{\partial \Phi^{(r)}(Q,\hat{Q})}{\partial r}\right\} \\
&= \operatorname{extr}\left\{\lim_{r\to 0}\frac{\partial}{\partial r}\left(-\operatorname{Tr}[Q\hat{Q}]\right)\right\} + \operatorname{extr}\left\{\lim_{r\to 0}\frac{\partial}{\partial r}\left(\log \Psi_{\text{w}}^{(r)}(\hat{Q})\right)\right\} + \alpha \operatorname{extr}\left\{\lim_{r\to 0}\frac{\partial}{\partial r}\left(\log \Psi_{\text{out}}^{(r)}(\hat{Q})\right)\right\}
\end{aligned} \tag{92}$$

考虑 $r \to 0$ 以及 $\hat{Q}^0 = 0$，最终得到自由能的复本对称解

$$\Phi_{\text{rs}}(\alpha) = \operatorname{extr}_{Q,\hat{Q},q,\hat{q},m,\hat{m}}\left\{-m\hat{m} + \frac{1}{2}Q\hat{Q} + \frac{1}{2}q\hat{q} + \Psi_{\mathbf{w}}(\hat{Q},\hat{m},\hat{q}) + \alpha\Psi_{\text{out}}(Q,m,q;\rho_{\text{w}^\star})\right\} \tag{93}$$

其中,

$$\begin{aligned}
\Psi_{\text{w}}(\hat{Q},\hat{m},\hat{q}) &\equiv \mathbb{E}_{\xi}\left[\mathcal{Z}_{\text{w}^\star}\left(\hat{m}\hat{q}^{-1/2}\xi, \hat{m}\hat{q}^{-1}\hat{m}\right)\log \mathcal{Z}_{\text{w}}\left(\hat{q}^{1/2}\xi, \hat{Q}+\hat{q}\right)\right] \\
\Psi_{\text{out}}(Q,m,q;\rho_{\text{w}^\star}) &\equiv \mathbb{E}_{y,\xi}\left[\mathcal{Z}_{\text{out}^\star}\left(y, mq^{-1/2}\xi, \rho_{\text{w}^\star} - mq^{-1}m\right)\log \mathcal{Z}_{\text{out}}\left(y, q^{1/2}\xi, Q-q\right)\right] \\
\rho_{\mathbf{w}^\star} &= \lim_{n\to\infty}\mathbb{E}_{\mathbf{w}^\star}\frac{1}{n}\|\mathbf{w}^\star\|_2^2
\end{aligned} \tag{94}$$

在贝叶斯最优中，有

$$\mathcal{Z}_{\text{out}} = \mathcal{Z}_{\text{out}^\star} \qquad \mathcal{Z}_{\mathbf{w}} = \mathcal{Z}_{\mathbf{w}^\star} \tag{95}$$

并且根据 Nishimori conditions 有

$$Q = \rho_{\text{w}^\star}, \quad m = q = q, \quad \hat{Q} = 0, \quad \hat{m} = \hat{q} = \hat{q} \tag{96}$$

式 (93) 简化为

$$\Phi(\alpha)=\operatorname{extr}_{q,\hat{q}}\left\{-\frac{1}{2}q\hat{q}+\Psi_{\mathrm{w}}^{\mathrm{b}}\left(\hat{q}\right)+\alpha\Psi_{\mathrm{out}}^{\mathrm{b}}\left(q;\rho_{\mathrm{w}^\star}\right)\right\} \tag{97}$$

其中，

$$\begin{aligned}\Psi_{\mathrm{w}}(\hat{q})&=\mathbb{E}_{\xi}\left[\mathcal{Z}_{\mathrm{w}^\star}\left(\hat{q}^{1/2}\xi,\hat{q}\right)\log\mathcal{Z}_{\mathrm{w}^\star}\left(\hat{q}^{1/2}\xi,\hat{q}\right)\right],\\ \Psi_{\mathrm{out}}\left(q;\rho_{\mathrm{w}^\star}\right)&=\mathbb{E}_{y,\xi}\left[\mathcal{Z}_{\mathrm{out}\,\star}\left(y,q^{1/2}\xi,\rho_{\mathrm{w}^\star}-q\right)\log\mathcal{Z}_{\mathrm{out}\,\star}\left(y,q^{1/2}\xi,\rho_{\mathrm{w}^\star}-q\right)\right]\end{aligned} \tag{98}$$

通过

$$\begin{aligned}\hat{Q}&=-2\alpha\partial_{Q}\Psi_{\mathrm{out}}\,, && Q=-2\partial_{\hat{Q}}\Psi_{\mathrm{w}}\\ \hat{q}&=-2\alpha\partial_{q}\Psi_{\mathrm{out}}\,, && q=-2\partial_{\hat{q}}\Psi_{\mathrm{w}}\\ \hat{m}&=\alpha\partial_{m}\Psi_{\mathrm{out}}\,, && m=\partial_{\hat{m}}\Psi_{\mathrm{w}}\end{aligned} \tag{99}$$

可以分别求得各个序参量，但是根据 Nishimori conditions，我们只需要求出

$$\begin{aligned}\hat{q}&=\alpha\mathbb{E}_{y,\xi}\left[\mathcal{Z}_{\mathrm{out}^\star}\left(y,q^{1/2}\xi,\rho_{\mathrm{w}^\star}-q\right)f_{\mathrm{out}^\star}\left(y,q^{1/2}\xi,\rho_{\mathrm{w}^\star}-q\right)^{2}\right]\\ q&=\mathbb{E}_{\xi}\left[\mathcal{Z}_{\mathrm{w}^\star}\left(\hat{q}^{1/2}\xi,\hat{q}\right)f_{\mathrm{w}^\star}\left(\hat{q}^{1/2}\xi,\hat{q}\right)^{2}\right]\end{aligned} \tag{100}$$

为了给出 $\mathcal{Z}_{\mathrm{out}}$ 的解析形式，我们用一个高斯分布取极限的形式来代替式 (3)

$$P(y\mid z)=\frac{1}{\sqrt{2\pi\Delta}}\exp\left(\frac{y-\operatorname{sign}(z)}{2\Delta}\right),\Delta\to\infty \tag{101}$$

因此

$$\begin{aligned}\mathcal{Z}_{\mathrm{out}}\left(y,\omega,V\right)=&\int dz\frac{1}{\sqrt{2\pi\Delta}}\exp\left(\frac{y-\operatorname{sign}(z)}{2\Delta}\right)\frac{1}{\sqrt{2\pi V}}\exp\left(-\frac{(z-\omega)^{2}}{2V}\right)\\ =&\int_{0}^{\infty}dz\frac{1}{\sqrt{2\pi\Delta}}\exp\left(\frac{y-1}{2\Delta}\right)\frac{1}{\sqrt{2\pi V}}\exp\left(-\frac{(z-\omega)^{2}}{2V}\right)\\ &+\int_{-\infty}^{0}dz\frac{1}{\sqrt{2\pi\Delta}}\exp\left(\frac{y+1}{2\Delta}\right)\frac{1}{\sqrt{2\pi V}}\exp\left(-\frac{(z-\omega)^{2}}{2V}\right)\\ =&\mathcal{N}_{y}\left(1,\Delta^{*}\right)\left(1+\operatorname{erf}\left(\frac{\omega}{\sqrt{2V}}\right)\right)+\mathcal{N}_{y}\left(-1,\Delta^{*}\right)\frac{1}{2}\left(1-\operatorname{erf}\left(\frac{\omega}{\sqrt{2V}}\right)\right)\\ f_{\mathrm{out}\,\star}(y,\omega,V)=&\frac{\mathcal{N}_{y}\left(1,\Delta^{\star}\right)-\mathcal{N}_{y}\left(-1,\Delta^{\star}\right)}{\mathcal{Z}_{\mathrm{out}\,\star}(y,\omega,V)}\mathcal{N}_{\omega}(0,V)\end{aligned} \tag{102}$$

同样的，在 $P_{\mathbf{w}^\star}\sim\mathcal{N}(0,1)$ 时有

$$\mathcal{Z}_{\mathrm{w}^\star}(\gamma,\Lambda)=\frac{e^{\frac{\gamma^{2}}{2(\Lambda+1)}}}{\sqrt{\Lambda+1}},\quad f_{\mathrm{w}^\star}(\gamma,\Lambda)=\frac{\gamma}{1+\Lambda},\quad \partial_{\gamma}f_{\mathrm{w}^\star}(\gamma,\Lambda)=\frac{1}{1+\Lambda} \tag{103}$$

因此

$$q=\frac{\hat{q}}{1+\hat{q}} \tag{104}$$

$$\begin{aligned}\hat{q}&=\alpha\mathbb{E}_{y,\xi}\left[\mathcal{Z}_{\mathrm{out}\,\star}\left(y,q^{1/2}\xi,1-q\right)f_{\mathrm{out}\,\star}\left(y,q^{1/2}\xi,1-q\right)^{2}\right]\\ &=\frac{2}{\pi}\frac{\alpha}{1-q}\int D\xi\frac{e^{-\frac{q_{\mathrm{b}}\xi^{2}}{1-q}}}{\left(1+\operatorname{erf}\left(\frac{\sqrt{q}\xi}{\sqrt{2(1-q)}}\right)\right)}\end{aligned} \tag{105}$$

在 $\alpha \to \infty$ 时，有 $q \to 1$，在此极限下有

$$\int D\xi \frac{e^{-\frac{q_{\mathrm{b}}\xi^2}{1-q_{\mathrm{b}}}}}{\left(1+\operatorname{erf}\left(\frac{\sqrt{q_{\mathrm{b}}}\xi}{\sqrt{2(1-q_{\mathrm{b}})}}\right)\right)} = \int \mathrm{d}\xi \frac{\frac{-e^{\frac{\xi^2(q_{\mathrm{b}}+1)}{2(1-q_{\mathrm{b}})}}}{\sqrt{2\pi}}}{\left(1+\operatorname{erf}\left(\frac{\sqrt{q_{\mathrm{b}}}\xi}{\sqrt{2(1-q_{\mathrm{b}})}}\right)\right)} \simeq \int \mathrm{d}\xi \frac{\frac{-e^{\frac{\xi^2}{1-q_{\mathrm{b}}}}}{\sqrt{2\pi}}}{\left(1+\operatorname{erf}\left(\frac{\xi}{\sqrt{2(1-q_{\mathrm{b}})}}\right)\right)}$$
$$= \frac{\sqrt{1-q_{\mathrm{b}}}}{\sqrt{2\pi}} \int \mathrm{d}\eta \frac{e^{-\eta^2}}{1+\operatorname{erf}\left(\frac{\eta}{\sqrt{2}}\right)} = \frac{c_0}{\sqrt{2\pi}}\sqrt{1-q_{\mathrm{b}}} \tag{106}$$

其中，

$$c_0 \equiv \int \mathrm{d}\eta \frac{e^{-\eta^2}}{1+\operatorname{erf}\left(\frac{\eta}{\sqrt{2}}\right)} \simeq 2.83748 \tag{107}$$

代入式 (105) 有

$$\hat{q}_{\mathrm{b}} = k\frac{\alpha}{\sqrt{1-q_{\mathrm{b}}}} \tag{108}$$

其中 $k \equiv \frac{2c_0}{\pi\sqrt{2\pi}} \simeq 0.720647$。解得

$$q_{\mathrm{b}} = \frac{1}{2}\left(\alpha k\sqrt{\alpha^2k^2+4} - \alpha^2k^2\right) \underset{\alpha\to\infty}{\simeq} 1 - \frac{1}{\alpha^2k^2}, \quad \hat{q}_{\mathrm{b}} = k^2\alpha^2 \tag{109}$$

代入泛化误差的表达式 (10) 得到泛化误差在大 α 时的渐进行为

$$e_{\mathrm{g}}^{\mathrm{bayes}}(\alpha) = \frac{1}{\pi}\operatorname{acos}\left(\sqrt{q_{\mathrm{b}}}\right) \underset{\alpha\to\infty}{\simeq} \frac{1}{k\pi}\frac{1}{\alpha} \simeq \frac{0.4417}{\alpha} \tag{110}$$

5 Programming and simulation

5.1 GAMP

设置超参数

```
alpha = np.arange(0.01, 5.5, 0.5)  # alpha==M/N
rho = 1                            # rho 是稀疏度，权重不为0所占比例
N = 1000                           # M是数据量，N是模型维度
num_times = 10                     # 迭代次数
record = {"error": []}
```

GAMP 迭代过程

```
from Gaussian_integral import Gamp

for alf in alpha:
    print("当alpha为",alf,"时")
    M = int(N * alf + 1e-3)
    record["error"].append([])

    for _ in range(3):  # 独立重复实验3次
        # 老师模型
        #########################
        # 生成数据，scale是标准差
        F = np.random.normal(loc=0, scale=np.sqrt(1), size=(M, N))
        # 老师权重 ground_truth
        x_tm = np.random.randn(N, 1)
        y_tm = np.sign(np.sqrt(1/N) * np.dot(F,x_tm))

        # GAMP迭代
        #########################
        a, v, flag = Gamp(rho, M, N, num_times,y_tm, F)
        if flag:
            # 生成新数据
            F_new = np.random.normal(loc=0, scale=np.sqrt(1), size=(M, N))
            # 老师生成的标签
            y_tm_new = np.sign(np.sqrt(1/N) * np.dot(F_new, x_tm))
            # 学生得到的后验概率
            y_sm_new = np.sign(np.sqrt(1/N) * np.dot(F_new, a))
            error_gen = np.heaviside((-y_tm_new * y_sm_new), 1).mean()
            # 泛化误差
            logging.info(f"alf: {alf}, error: {error_gen:.5\%}")
            record["error"][-1].append(error_gen)
        else:
            logging.info("AMP fails")
```

绘图

```python
# 计算标准差和均值
ampErrorStd = np.std(ampError, axis=1)
ampErrorMean = np.mean(ampError, axis=1)
import matplotlib.pyplot as plt
# 绘制误差线图
plt.errorbar(alpha, ampErrorMean, ampErrorStd, fmt='ro',markerfacecolor='none',
                                          capsize=4, label='Gamp')
plt.xlabel("$\\alpha$")
plt.ylabel("Test err")
plt.legend()
plt.show()
```

其中用到的 GAMP 函数为

```python
def Gamp(rho, M, N, num_times, y_tm,F):
    from tqdm import tqdm

    # 初始化
    a = rho * np.ones((N,1)) * Gaussian_integral_mean(0, 1)
    v = rho * np.ones((N,1)) * Gaussian_integral_var(0, 1) - a**2
    gout = np.zeros_like(y_tm)
    pw_gout = np.zeros_like(y_tm)
    # 开始迭代

    ### 防止发散
    import sys
    epsilon = sys.float_info.epsilon

    flag =False
    for t in tqdm(range(num_times)):
        ### 引入中间变量
        V_ = 1/N * np.dot(F*F, v)
        w_ = np.sqrt(1/N) * np.dot(F, a) - V_*gout
        ### 更新gout
        for j in range(M):
            gout[j] = Indicator_integral_mean(w_[j], np.sqrt(V_[j]), y_tm[j]) \
                     /(Indicator_integral_PDF(w_[j], np.sqrt(V_[j]), y_tm[j]) \
                     + epsilon) - w_[j]
            ### 计算partial
            pw_gout[j] = (Indicator_integral_var(w_[j], np.sqrt(V_[j]), y_tm[j]) \
                         -2*w_[j]*Indicator_integral_mean(w_[j], np.sqrt(V_[j]),
                         y_tm[j]))\
                         /(Indicator_integral_PDF(w_[j], np.sqrt(V_[j]), y_tm[j])
                         + epsilon)\
                         + w_[j] * w_[j] - gout[j] * gout[j] -V_[j]
        gout = 1/V_ * gout
        pw_gout = 1/V_**2 * pw_gout
```

```python
        ### 引入中间变量
        Xi = -1/(1/N * np.dot((F*F).T, pw_gout))
        R = a + Xi * np.sqrt(1/N) *np.dot(F.T, gout)

        a_new = np.zeros_like(a)
        v_new = np.zeros_like(v)
        ### 更新a, v
        for i in range(N):
            a_new[i] = prior_integral_mean(R[i],np.sqrt(Xi)[i],0,1)\
                    /(prior_integral_PDF(R[i],np.sqrt(Xi)[i],0,1)+ epsilon)

            v_new[i] = prior_integral_var(R[i],np.sqrt(Xi)[i],0,1)\
                    /(prior_integral_PDF(R[i],np.sqrt(Xi)[i],0,1)+ epsilon)\
                    - a_new[i]*a_new[i]
        flag = True
        if np.max(np.abs(np.stack([a_new-a, v_new-v], axis=0))) < 3e-3:
            flag = True
            break
        a = a_new
        v = v_new
    return a, v, flag
```

里面使用了一些自定义的高斯积分

```python
# 高斯测度积分
#########################
import scipy.integrate as spi

# 高斯分布的积分
def Gaussian_integral_PDF(mu,sigma):
    ### sigma是标准差

    # 定义被积函数
    def integrand(x):
        return 1/(np.sqrt(2*np.pi)*sigma) *\
            np.exp(-(x-mu)**2/(2*sigma**2))

    # 调用 quad() 函数计算积分
    result, error = spi.quad(integrand, -np.inf, np.inf)
    return result

# 高斯均值的积分
def Gaussian_integral_mean(mu, sigma):
    ### sigma是标准差

    # 定义被积函数
    def integrand(x):
```

```
        return 1 / (np.sqrt(2 * np.pi) * sigma)* x * \
               np.exp(-(x - mu) ** 2 / (2 * sigma ** 2))

    # 调用 quad() 函数计算积分
    result, error = spi.quad(integrand, -np.inf, np.inf)
    return result

# 高斯方差的积分
def Gaussian_integral_var(mu, sigma):
    ### sigma是标准差

    # 定义被积函数
    def integrand(x):
        return 1 / (np.sqrt(2 * np.pi) * sigma) * x**2 * \
               np.exp(-(x - mu) ** 2 / (2 * sigma ** 2))

    # 调用 quad() 函数计算积分
    result, error = spi.quad(integrand, -np.inf, np.inf)
    return result


#########################
# 高斯测度的概率积分
#########################

# P_out测度积分
#########################
# P_out是indicator function的情况

# 高斯分布的积分
def Indicator_integral_PDF(mu,sigma, ref):
    ### sigma是标准差

    # 定义被积函数
    def integrand(x):
        return 1/(np.sqrt(2*np.pi)*sigma) * \
               np.where(np.sign(x)==ref,1,0) * \
               np.exp(-(x-mu)**2/(2*sigma**2))

    # 调用 quad() 函数计算积分
    result, error = spi.quad(integrand, -np.inf, np.inf)
    return result

# 高斯均值的积分
def Indicator_integral_mean(mu, sigma, ref):
    ### sigma是标准差
```

```
    # 定义被积函数
    def integrand(x):
        return 1 / (np.sqrt(2 * np.pi) * sigma)* x *\
               np.where(np.sign(x)==ref,1,0) * \
               np.exp(-(x - mu) ** 2 / (2 * sigma ** 2))

    # 调用 quad() 函数计算积分
    result, error = spi.quad(integrand, -np.inf, np.inf)
    return result

# 高斯方差的积分
def Indicator_integral_var(mu, sigma, ref):
    ### sigma是标准差

    # 定义被积函数
    def integrand(x):
        return 1 / (np.sqrt(2 * np.pi) * sigma) * x**2 *\
               np.where(np.sign(x)==ref,1,0) * \
               np.exp(-(x - mu) ** 2 / (2 * sigma ** 2))

    # 调用 quad() 函数计算积分
    result, error = spi.quad(integrand, -np.inf, np.inf)
    return result

# P_0测度积分
#########################
# P_0是随机高斯的情况N(MU, SIG^2)

# 高斯分布的积分
def prior_integral_PDF(mu,sigma, MU, SIG):
    ### sigma是标准差

    # 定义被积函数
    def integrand(x):
        return 1/(np.sqrt(2*np.pi)*sigma) * \
               1/(np.sqrt(2*np.pi)*SIG) * np.exp(-(x-MU)**2/(2*SIG**2)) * \
               np.exp(-(x-mu)**2/(2*sigma**2))

    # 调用 quad() 函数计算积分
    result, error = spi.quad(integrand, -np.inf, np.inf)
    return result

# 高斯均值的积分
def prior_integral_mean(mu, sigma, MU, SIG):
    ### sigma是标准差

    # 定义被积函数
```

```python
    def integrand(x):
        return 1 / (np.sqrt(2 * np.pi) * sigma)* x *\
                1/(np.sqrt(2*np.pi)*SIG) * np.exp(-(x-MU)**2/(2*SIG**2)) * \
                np.exp(-(x - mu) ** 2 / (2 * sigma ** 2))

    # 调用 quad() 函数计算积分
    result, error = spi.quad(integrand, -np.inf, np.inf)
    return result

# 高斯方差的积分
def prior_integral_var(mu, sigma, MU, SIG):
    ### sigma是标准差

    # 定义被积函数
    def integrand(x):
        return 1 / (np.sqrt(2 * np.pi) * sigma) * x**2 *\
                1/(np.sqrt(2*np.pi)*SIG) * np.exp(-(x-MU)**2/(2*SIG**2)) * \
                np.exp(-(x - mu) ** 2 / (2 * sigma ** 2))

    # 调用 quad() 函数计算积分
    result, error = spi.quad(integrand, -np.inf, np.inf)
    return result
```

5.2 SE

设置超参数

```python
import numpy as np

N = 1000  # M是数据量, N是模型维度
num_times = 20  # 迭代次数
alpha = np.arange(0.01, 10.1, 0.1)
```

迭代求解 SE

```python
from Gaussian_integral import erf2_integral_PDF
from Gaussian_integral import perception_integral_PDF

# 初始化
q = np.zeros_like(alpha)
qhat = np.zeros_like(alpha)

# 开始迭代
flag =False
from tqdm import tqdm
for t in tqdm(range(num_times)):
    q_new = np.zeros_like(alpha)
    qhat_new = np.zeros_like(alpha)
```

```python
    for i in range(len(alpha)):
        q_new[i] = qhat[i]/(1+qhat[i])
        qhat_new[i] = 2*alpha[i]/(np.pi*(1-q_new[i])) * perception_integral_PDF(
                                                     q_new[i])
        # print('q',q_new)
        # print('qhat', qhat_new)

    # flag = True
    if np.max(np.abs(np.stack([q_new - q, qhat_new - qhat], axis=0))) < 1e-3:
        flag = True
        break
    q = q_new
    qhat = qhat_new

error_gen = 1/np.pi*(np.arccos(np.sqrt(q)))
```

绘图

```python
import matplotlib.pyplot as plt

# 绘制误差线图
plt.plot(alpha, error_gen,'b',label='SE')
plt.plot(alpha, 0.4417/alpha,'c',label='asymptotic solution' )
plt.xlabel("$\\alpha$")
plt.ylabel("Test err")
plt.legend()
plt.show()
```

其中用到两个自定义的测度积分

```python
from scipy.special import erf
# erf是随机高斯的情况N(MU, SIG^2)

# 高斯分布的积分
def erf2_integral_PDF(mu,sigma,mu_erf, sigma_erf):
    ### sigma是标准差

    # 定义被积函数
    def integrand(x):
        return 1 / (np.sqrt(2 * np.pi) * sigma) * \
               1 / ((2 * np.pi) * sigma_erf**2) * \
               1 / (1+erf(x/(np.sqrt(2)*sigma_erf))+1e-9)* \
               np.exp(-(x-mu_erf) ** 2 / (sigma_erf ** 2)) * \
               np.exp(-(x - mu) ** 2 / (2 * sigma ** 2))

    # 调用 quad() 函数计算积分
    result, error = spi.quad(integrand, -np.inf, np.inf)
    return result
```

```python
def perception_integral_PDF(q):
    ### sigma是标准差

    # 定义被积函数
    def integrand(x):
        return 1 / np.sqrt(2 * np.pi) * \
               1 / (1+erf(x*np.sqrt(q)/np.sqrt(2*(1-q)))+1e-9)* \
               np.exp(- x** 2 *(q+1)/(2*(1-q)))

    # 调用 quad() 函数计算积分
    result, error = spi.quad(integrand, -np.inf, np.inf)
    return result
```

Reference

[1] ZDEBOROVÁ L KRZAKALA F. Statistical physics of inference: thresholds and algorithms. Advances in Physics, 2016, 65(5): 453-552. eprint: https://doi.org/10.1080/00018732.2016.121139. https://doi.org/10.1080/00018732.2016.1211393. DOI: 10.1080/00018732.2016.1211393.

[2] AUBIN B, KRZAKALA F, LU Y, Generalization error in high-dimensional perceptrons: Approaching bayes error with convex optimization. Advances in Neural Information Processing Systems, 2020, 33: 12199-12210.

[3] BARBIER J, KRZAKALA F, MACRIS N, Optimal errors and phase transitions in high-dimensional generalized linear models. Proceedings of the National Academy of Sciences, 2019, 116(12): 5451-5460.

[4] LUCIBELLO C, PITTORINO F, PERUGINI G, Deep learning via message passing algorithms based on belief propagation. Machine Learning: Science and Technology, 2022, 3(3): 035005.

[5] GABRIÉ M. Mean-field inference methods for neural networks. Journal of Physics A: Mathematical and Theoretical, 2020, 53(22): 223002.

[6] AUBIN B, MAILLARD A, KRZAKALA F, The committee machine: Computational to statistical gaps in learning a two-layers neural network. Advances in Neural Information Processing Systems, 2018, 31.

[7] RANGAN S. Generalized approximate message passing for estimation with random linear mixing//2011 IEEE International Symposium on Information Theory Proceedings. 2011: 2168-2172.

[8] CORNACCHIA E, MIGNACCO F, VEIGA R, Learning curves for the multi-class teacher-student perceptron. Machine Learning: Science and Technology, 2022.

[9] LOUREIRO B, SICURO G, GERBELOT C, Learning gaussian mixtures with generalized linear models: Precise asymptotics in high-dimensions. Advances in Neural Information Processing Systems, 2021, 34: 10144-10157.

[10] DONOHO D L, MALEKI A, MONTANARI A. Message-passing algorithms for compressed sensing. Proceedings of the National Academy of Sciences, 2009, 106(45): 18914-18919.

[11] ADVANI M GANGULI S. An equivalence between high dimensional Bayes optimal inference and M-estimation. Advances in Neural Information Processing Systems, 2016, 29.

[12] PARK K, KWON C, PARK Y. One-step replica-symmetry-breaking solution for a perceptron learning with weight mismatch. Journal of Physics A: Mathematical and General, 1996, 29(7): 1397.

[13] BARBIER J, KRZAKALA F, MACRIS N, Phase transitions, optimal errors and optimality of message-passing in generalized linear models. 2017.

[14] KRZAKALA F, MÉZARD M, SAUSSET F, Probabilistic reconstruction in compressed sensing: algorithms, phase diagrams, and threshold achieving matrices. Journal of Statistical Mechanics: Theory and Experiment, 2012, 2012(08): P08009.

[15] ADVANI M, LAHIRI S, GANGULI S. Statistical mechanics of complex neural systems and high dimensional data. Journal of Statistical Mechanics: Theory and Experiment, 2013, 2013(03): P03014.

[16] ZDEBOROVÁ L KRZAKALA F. Statistical physics of inference: Thresholds and algorithms. Advances in Physics, 2016, 65(5): 453-552.

[17] BAYATI M MONTANARI A. The dynamics of message passing on dense graphs, with applications to compressed sensing. IEEE Transactions on Information Theory, 2011, 57(2): 764-785.

[18] KABASHIMA Y, KRZAKALA F, MÉZARD M, Phase Transitions and Sample Complexity in Bayes-Optimal Matrix Factorization. IEEE Transactions on Information Theory, 2016, 62(7): 4228-4265. DOI: 10.1109/TIT.2016.2556702.