STSCI 6940:

Lecture 3 Date: 02/16

rank-one spiked Wigner model.

$$Y_{ij} = \sqrt{\frac{\lambda}{n}}\, x_i x_j + W_{ij}$$

where $x_i \overset{iid}{\sim} P_0$ zero mean, unit variance.

$$W_{ij} \sim N(0,1) \text{ iid for } i < j.$$
$$W_{ij} = W_{ji}$$

[Let's ignore the diagonal terms].

$$Y = \sqrt{\frac{\lambda}{n}}\, x x^T + W$$

The posterior measure:

$$dP(x \mid Y) \propto e^{\underbrace{\sum_{i<j} \sqrt{\frac{\lambda}{n}} Y_{ij} x_i x_j - \frac{\lambda}{2n} x_i^2 x_j^2}_{H_n(x)}} dP_0^n(x)$$

Free energy:

$$F_n(\lambda) = \mathbb{E} \log \int e^{H_n(x)} \, dP_0^n(x).$$

Limit of $\frac{1}{n} F_n(\lambda)$?

Theorem (replica-symmetric formula for F_n):

$$\forall \lambda \geq 0: \quad \lim_{n \to +\infty} \frac{1}{n} F_n(\lambda) = \sup_{q \geq 0} \left\{ \psi(\lambda q) - \frac{\lambda q^2}{4} \right\}.$$

$$\psi(T) = \underset{x_0 \sim P_0,\ z \sim N(0,1)}{\mathbb{E}} \log \int e^{\sqrt{T} z x + T x x_0 - \frac{T}{2} x^2} dP_0(x).$$

ψ is the free energy of a scalar Gaussian

additive model ; $y = \sqrt{T}\, x_0 + z$ $z \sim N(0,1)$, $x_0 \sim P_0$.

- physics prediction :
  - Lesieur - Krzakala - Zdeborová 2015
  - Kabashima . 2008 - 2009 .
- rigorous proof :
  - Deshpande, Abbé, Montanari (2014) AMP.
  - Lelarge, Miolane 2017

(Guerra / interpolation method + Cavity method)

- Heuristic derivation :

Cavity method.

Main observation : (a_n) a sequence of reals,

such that $\lim_{n\to+\infty} a_n = \ell$.

$$\Rightarrow \lim_{n\to+\infty} \frac{1}{n} \sum_{i=1}^{n} a_i = \ell .$$

Apply this to $a_n = F_{n+1}(\lambda) - F_n(\lambda)$.

$$F_{n+1}(\lambda) - F_n(\lambda) = \mathbb{E} \log \frac{Z_{n+1}(\lambda)}{Z_n(\lambda)} ,$$

$$H_{n+1}(x) = \sum_{1\le i<j\le n+1} \sqrt{\frac{\lambda}{n+1}}\, w_{ij}\, x_i x_j + \frac{\lambda}{n+1} x_i x_j x_{0i} x_{0j} - \frac{\lambda}{2(n+1)} x_i^2 x_j^2 .$$

$$Y_{ij} = \sqrt{\frac{\lambda}{n+1}}\, x_{0i} x_{0j} + w_{ij}.$$

coupling Z_n and Z_{n+1} with the same

$$(w_{ij})_{1\le i<j\le n} \quad \text{and} \quad (x_{0i})_{1\le i\le n}$$

$$x = (x_i)_{i=1}^{n}, \quad x_{n+1} = \varepsilon .$$

$$H_{n+1}(x,\varepsilon) = \sum_{1\le i<j\le n} \underset{\underset{\frac{\lambda_n}{n}}{\shortparallel}}{\sqrt{\frac{\lambda}{n+1}}}\, w_{ij}\, x_i x_j + \frac{\lambda}{n+1} x_i x_j x_{0i} x_{0j} - \frac{\lambda}{2(n+1)} x_i^2 x_j^2$$

$$h_{n+1}(x,\varepsilon) \left\{ + \sum_{i=1}^{n+1} \sqrt{\frac{\lambda}{n+1}}\, w_{i\,n+1}\, x_i\, \varepsilon + \frac{\lambda}{n+1} x_i\, x_{0i} \cdot \varepsilon_0\, \varepsilon - \frac{\lambda}{2(n+1)} \cdot x_i^2 \cdot \varepsilon^2 \right.$$

$$\tilde{H}_{n+1}(x,\varepsilon) = H_n(x) + h_{n+1}(x,\varepsilon) .$$

Let's write :

$$\frac{Z_{n+1}(\lambda)}{Z_n(\lambda)} = \frac{\int e^{H_{n+1}(x,\varepsilon)}\, dP_0(\varepsilon)\, dP_0^n(x)}{\int e^{H_n(x)}\, dP_0^n(x)} .$$

$$= \underbrace{\frac{\int e^{\tilde{H}_{n+1}(x,\varepsilon)}\, dP_0(\varepsilon)\, dP_0^n(x)}{\int e^{H_n(x)}\, dP_0^n(x)}}_{A} \cdot \underbrace{\frac{\int e^{H_{n+1}(x,\varepsilon)}\, dP_0(\varepsilon)\, dP_0^n(x)}{\int e^{\tilde{H}_{n+1}(x,\varepsilon)}}}_{B}$$

$$\lambda_n = \lambda \frac{n}{n+1} = \lambda \left(1 - \frac{1}{n+1}\right).$$

$$\mathbb{E} \log B = F_{n+1}(\lambda_n) - F_{n+1}(\lambda).$$

$$= F'_{n+1}(\lambda) \cdot (\lambda_n - \lambda) + o_n(1).$$

(hopefully).

$$= F'_{n+1}(\lambda) \left(- \frac{\lambda}{n+1} \right).$$

(I - MMSE relation)

$$F'_{n+1}(\lambda) = \frac{n+1}{4} \mathbb{E} \left[\left\langle \left(\frac{1}{n} \sum_{i=1}^{n} x_i x_{0i} \right)^2 \right\rangle_{n,\lambda} \right]$$

$$\langle \cdot \rangle_{n+1,\lambda} := \frac{\int \cdot \, e^{H_{n+1}(x)} \, dP_0^{n+1}(x)}{\int e^{-} \_\_\_}.$$

$$\mathbb{E} \log B = -\frac{\lambda}{4} \mathbb{E} \left\langle \Big(\underbrace{\frac{1}{n} \sum_{i=1}^{n} x_i x_{0i}} \Big)^2 \right\rangle_{n+1,\lambda}.$$

$$R_{1,0} = \frac{1}{n} \sum_{i=1}^{n} x_i \cdot x_{0i} = x \cdot x_0,$$

Assumption (Replica symmetry):

$$x \sim \mathbb{P}(\cdot \mid Y).$$

$$x \cdot x_0 \xrightarrow[n \to +\infty]{} q$$

- Nishimori identities:

$$x_1, x_2 \sim \mathbb{P}(\cdot \mid Y) \text{ independently.}$$

claim : $(X_1, X_2, Y) \stackrel{d}{=} (X_0, X_1, Y)$.

pf : $\psi : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{n \times n}$

$$\mathbb{E}[\psi(X_0, X_1, Y)] = \mathbb{E}\big[\mathbb{E}[\psi(X_0, X_1, Y) \mid Y]\big].$$

$$= \mathbb{E}[\psi(X_2, X_1, Y)]$$

This generalizes to $(X_0, X_1, \dots, X_n, Y)$

$$\stackrel{d}{=} (X_1, X_2, \dots, X_{n+1}, Y).$$

Experiment :

- $X_0 \sim P_0$.
- generate Y from the model.
- $X_1, X_2 \sim \mathbb{P}(\cdot \mid Y)$ independently.

Claim :

$$(X_1, X_2, Y) \stackrel{d}{=} (X_0, X_1, Y).$$

$$\Rightarrow \quad (X_1, X_0) \stackrel{d}{=} (X_1, X_2).$$

replica symmetry implies :

$$X_1 \cdot X_2 \longrightarrow q.$$

Assuming RS : $\boxed{\mathbb{E} \log B \longrightarrow -\frac{\lambda}{4} q^2.}$

$$A = \frac{\int e^{\tilde{H}_{n+1}(x,\varepsilon)} \, dP_0(\varepsilon)\, dP_0^n(x)}{\int e^{H_n(x)} \, dP_0^n(x)}$$

$$= \frac{\int e^{H_n(x)} \left(\int e^{h_{n+1}(x,\varepsilon)} \, dP_0(\varepsilon) \right) dP_0^n(x)}{\int e^{H_n(x)} \, dP_0^n(x)} .$$

$$= \left\langle \int e^{h_{n+1}(x,\varepsilon)} \, dP_0(\varepsilon) \right\rangle_{n,\lambda} .$$

$$h_{n+1}(x,\varepsilon) = \sqrt{\frac{\lambda}{n+1}} \left(\sum_{i=1}^{n} g_i \; x_i \right) \varepsilon$$
$$+ \frac{\lambda}{n+1} \left(\sum_{i=1}^{n} x_i x_{0i} \right) \varepsilon \varepsilon_0$$
$$- \frac{\lambda}{2(n+1)} \left(\sum_{i=1}^{n} x_i^2 \right) \varepsilon^2 .$$

$x \sim \mathbb{P}_n(\cdot \mid Y)$

$$G_n(x) = \frac{1}{\sqrt{n+1}} \sum_{i=1}^{n} g_i \, x_i$$

$(G_n(x))_{x \in \mathbb{R}^n}$ centered Gaussian process.

$$\mathbb{E}\left[G_n(x) \cdot G_n(x') \right] = \frac{1}{n+1} \sum_{i=1}^{n} x_i x_i'$$

$$= \frac{n}{n+1} \; x \cdot x' .$$

leap of faith.

only thing that matters in the covariance structure are $x, x' \sim \mathbb{P}(\cdot \mid Y)$ independently.

in which case.

$$x \cdot x' \longrightarrow q \quad (\text{RS}).$$

$(G_n(x))$ is equivalent to a Gaussian process

$$\sum_{i=1}^{n} \sqrt{q}\, z_i \cdot x_i$$

- $\mathbb{E} \log A \simeq \mathbb{E} \log \int e^{\sqrt{\lambda q}\, z \varepsilon + \lambda q \varepsilon \varepsilon_0 - \frac{\lambda q^2}{2}} dP_0(\varepsilon)$.

$$= \psi(\lambda q).$$

Conclusion: $F_{n+1}(\lambda) - F_n(\lambda) \longrightarrow \psi(\lambda q) - \lambda q^2/4$.

$$\Rightarrow \quad \frac{1}{n} F_n(\lambda) \longrightarrow \psi(\lambda q) - \lambda q^2/4 .$$

for q such that $x \cdot x_0 \to q$ when $x \sim \mathbb{P}(\cdot \mid Y)$.

Heuristic provides
Additional information; sup is acheived at q such.

Main idea: Compare

$$① \quad Y_{ij} = \sqrt{\tfrac{\lambda}{n}}\, x_i x_j + w_{ij}.$$

to

$$② \quad y_i = \sqrt{\lambda q}\, x_i + z_i \quad \text{indep } 1 \le i \le n$$

$\sim N(0,1)$.

Interpolation argument:
Smoothly interpolate between ① and ②

$$\begin{cases} Y_{ij} = \sqrt{\frac{t\lambda}{n}}\, x_i x_j + w_{ij}. \\ y_i = \sqrt{(1-t)\lambda q}\, x_i + z_i \end{cases}$$