Lecture 20: Approximate Message Passing

State evolution: $\forall$ nice $\Psi : \mathbb{R}^2 \to \mathbb{R}$, $t \in \mathbb{N}$,

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \Psi(\underset{\substack{\text{t}^{th} \text{ iterate} \\ \text{of AMP}}}{x_t^i}, \underset{\text{signal}}{X^i}) = \underset{\substack{x \sim \{\pm 1\} \\ g \sim N(0,1)}}{\mathbb{E}} \left[\Psi(\mu_t x + \sigma_t g, x) \right]$$

$$\mu_t = \sqrt{\lambda} \sigma_t^2, \qquad \sigma_t^2 = \gamma_t / \lambda, \qquad \gamma_{t+1} = \lambda(1 - \text{mmse}(\gamma_t))$$

Lemma 1: $\quad \text{MSE}_{\text{AMP}}(t; \lambda) = 1 - \dfrac{\gamma_{t+1}^2}{\lambda^2}$.

Pf:

$$\text{MSE}_{\text{AMP}}(t; \lambda, n) = \frac{1}{n^2} \mathbb{E}\left[\left\| XX^T - \hat{x}^t (\hat{x}^t)^T \right\|_F^2 \right]$$

$$= \underbrace{\frac{\mathbb{E}\left[\|X\|_2^4 \right]}{n^2}}_{1} - 2 \cdot \underbrace{\frac{\mathbb{E}\left(\langle \hat{x}^t, X \rangle^2 \right)}{n^2}}_{\text{I}} + \underbrace{\frac{\mathbb{E}\|x^t\|_2^4}{n^2}}_{\text{II}}$$

for ⓘ, we can use state evolution.

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} X_i (\hat{x}^t)_i = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} X_i f_t(x_i^t)$$

$$= \underset{X, \zeta}{\mathbb{E}}\left[X f_t \left(\mu_t X + \sigma_t \zeta \right) \right]$$

$$= \mu_{t+1} / \sqrt{\lambda} \qquad \text{(by defn of } \mu_{t+1})$$

$$= \gamma_{t+1} / \lambda,$$

So $\quad \lim_{n \to \infty} \dfrac{1}{n^2} \mathbb{E}\left[\langle \hat{x}^t, X \rangle^2 \right] = \dfrac{\gamma_{t+1}^2}{\lambda^2}$

Similar calculation for ②:

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} (\hat{x}_t)_i^2 = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} f_t (X_i^t)^2$$

$$= \mathbb{E} \left[\mathbb{E}(X \mid \mu_t X + \sigma_t Z)^2 \right]$$

$\mathbb{E}(xx^T \mid y)$

$= \mathbb{E}(x \mid y x_{t+1})^{\otimes 2}$

$$= \sigma_{t+1}^2 \qquad (\text{by defn } \oint \sigma_{t+1}^2)$$

$$= \gamma_{t+1} / \lambda ,$$

So $\quad \lim_{n \to \infty} \frac{1}{n^2} \mathbb{E}\left(\| \hat{x}^t \|_2^4 \right) = \frac{\gamma_{t+1}^2}{\lambda^2}.$ $\qquad \square$

$\underline{\text{Lemma 2}}$: $\quad MMSE(\lambda) = \lim_{t \to \infty} MSE_{AMP}(t; \lambda)$

$\underline{Pf}$: We will use the following:

Fact (I-MMSE relation):

$Y(\lambda) = \sqrt{\frac{\lambda}{n}} XX^T + W$

$$\frac{1}{n} \cdot \frac{\partial}{\partial \lambda} I\left(XX^T ; Y(\lambda) \right) = \frac{1}{4} MMSE(\lambda, n)$$

Mutual information between signal and observation

Note

when $\lambda = 0$, no information about signal present

when $\lambda = \infty$, there is full information about signal

$$\lim_{n \to \infty} \frac{1}{n} I(XX^T; Y(0)) = 0$$

$$\lim_{\lambda \to \infty} \lim_{n \to \infty} \frac{1}{n} I(XX^T; Y(\lambda)) = \lg 2, \quad \text{so}$$

$$\lg 2 = \lim_{\lambda \to \infty} \lim_{n \to \infty} \frac{1}{n} \left[I\left(XX^T; Y(\lambda)\right) - I(XX^T; Y(0)) \right]$$

(I-MMSE)
$$= \lim_{n \to \infty} \frac{1}{4} \int_0^\infty MMSE(\lambda, n) \, d\lambda$$

(Bayes optimality)
$$\leq \lim_{t \to \infty} \limsup_{n \to \infty} \frac{1}{4} \int_0^\infty MSE_{AMP}(t; \lambda, n) \, d\lambda$$

(dominated convergence thm).
$$= \frac{1}{4} \int_0^\infty \left(1 - \frac{\gamma_\bullet(\lambda)^2}{\lambda^2} \right) d\lambda$$

If can show

$$\frac{1}{4} \int_0^\infty \left(1 - \frac{\gamma_\bullet(\lambda)^2}{\lambda^2} \right) d\lambda = \lg 2,$$

then $\leq$ above is an equality, so AMP is

Bayes-optimal as desired.

(we will show this is the antiderivative of MSE_{AMP} w.r.t. λ)

Define $\psi(\gamma, \lambda) \doteq \frac{\lambda}{4} + \frac{\gamma^2}{4\lambda} - \frac{\gamma}{2} + I(\gamma),$

where $I(\gamma) \triangleq I(X; \sqrt{\gamma} X + 3)$.

Note that by design,

$$\frac{\partial}{\partial \lambda} \Psi(\gamma, \lambda) = \frac{1}{4}\left(1 - \frac{\gamma^2}{\lambda^2}\right)$$

$$\frac{\partial}{\partial \gamma} \Psi(\gamma, \lambda) = \frac{\gamma}{2\lambda} - \frac{1}{2} + \frac{1}{2} \text{mmse}(\gamma)$$

$$= 0 \quad \text{if} \quad \gamma = \gamma_*(\lambda).$$

so $$\underline{\underline{\frac{\partial}{\partial \lambda} \Psi(\gamma_*(\lambda), \lambda) = \frac{1}{4}\left(1 - \frac{\gamma_*(\lambda)^2}{\lambda^2}\right)}}$$

Furthermore, $\lim\limits_{\lambda \to 0} \Psi\left(\gamma_*(\lambda), \lambda\right) = 0$ as

$\gamma_*(\lambda) \leq \lambda \to 0$ and $I(\gamma_*(\lambda)) \to 0$, so

$$\frac{1}{4} \int_0^\infty \left(1 - \frac{\gamma_*(\lambda)^2}{\lambda^2}\right) = \lim\limits_{\lambda \to \infty} \Psi\left(\gamma_*(\lambda), \lambda\right)$$

$$\gamma^*(\lambda) \to \lambda \quad \text{as} \quad \lambda \to \infty, \text{ so}$$

$$\frac{\lambda}{4} + \frac{\gamma^*(\lambda)^2}{4\lambda} - \frac{\gamma^*(\lambda)}{2} \to 0.$$

Furthermore, $I(\gamma^0(\lambda)) \to \lg 2$, so
the proof is complete. $\qquad\qquad\qquad$ □

Proof of I-MMSE relation:

Let us consider a more general setting where
X is some random vector in $\mathbb{R}^N$, and
we observe $Y = \sqrt{\lambda} X + g$ for $g \sim N(0, I_d)$.

$$MMSE(\lambda) \triangleq \mathbb{E}\left[\|X - \mathbb{E}(X|Y)\|^2\right] \quad (\text{unnormalized}).$$

Mutual information

$$I(X; Y) \triangleq KL\left(\underbrace{P_{X,Y}}_{\text{Joint dist. of } X, Y} \| \underbrace{P_X \otimes P_y}_{\substack{\text{product of} \\ \text{marginal dist.'s}}}\right)$$

$$= \mathbb{E}_{X,Y}\left[\lg \frac{dP_{X,Y}}{dP_X \otimes P_y}(X,Y)\right]$$

$$= \mathbb{E}_{X,Y}\left[\lg\left\{\frac{\exp\left(-\frac{1}{2}\|Y - \sqrt{\lambda} X\|^2\right)}{\int dP_X(x) \exp\left(-\frac{1}{2}\|Y - \sqrt{\lambda} x\|^2\right)}\right\}\right]$$

$$= -\mathop{\mathbb{E}}_{X,Y} \lg \int dP_X(x) \exp\left(\sqrt{\lambda}\langle x, Y\rangle - \underline{\sqrt{\lambda}\langle X, Y\rangle} - \tfrac{1}{2}\|x\|^2 + \underline{\tfrac{\lambda}{2}\|X\|^2} \right)$$

$$= \underline{\tfrac{\lambda}{2}\mathbb{E}\|X\|^2} - \underbrace{\mathop{\mathbb{E}}_{X,Y} \lg \overbrace{\int dP_X(x) \exp\left(\sqrt{\lambda}\langle x, Y\rangle - \tfrac{1}{2}\|x\|^2 \right)}^{\substack{\text{partition function} \\ \text{for Gibbs measure } \mu = \Pr[X=\cdot\,|Y]}}}_{\triangleq F(\lambda) \quad (\text{free energy})}$$

w.t.s $\quad \dfrac{\partial}{\partial \lambda} F(\lambda) = \tfrac{1}{2}\left(\mathbb{E}\|X\|^2 - \text{MMSE}(\lambda) \right)$

$$\left(= \tfrac{1}{2}\mathbb{E}\|\mathbb{E}(X|Y)\|^2 \right)$$

$$\frac{d}{d\lambda} F(\lambda) = \mathop{\mathbb{E}}_{X,g} \mathop{\mathbb{E}}_{x\sim\mu} \left[\underbrace{\langle x, X\rangle}_{\mathbb{E}\|\mathbb{E}(X|Y)\|^2} + \underline{\tfrac{1}{2\sqrt{\lambda}}\langle x, g\rangle} - \tfrac{1}{2}\|x\|^2 \right]$$

$$\mathop{\mathbb{E}}_{X} \mathop{\mathbb{E}}_{g} \mathop{\mathbb{E}}_{x\sim\mu} \left[\tfrac{1}{2\sqrt{\lambda}}\langle x, g\rangle \right]$$

$$\left(\substack{\text{Gaussian} \\ \text{int. by} \\ \text{parts}}\right) = \mathop{\mathbb{E}}_{X} \mathop{\mathbb{E}}_{g} \text{div}_g \left\{ \mathop{\mathbb{E}}_{x\sim\mu}\left[\tfrac{1}{2\sqrt{\lambda}} x \right] \right\}$$

Note, for any $i \in [N]$,

$$\frac{\partial}{\partial g_i} \mathop{\mathbb{E}}_{x \sim \mu}[x_i] = \sqrt{\lambda}\left(\mathop{\mathbb{E}}_{x \sim \mu}[x_i^2] - \mathop{\mathbb{E}}_{x \sim \mu}[x_i]^2\right)$$

(see prop. below)

So $\mathop{\mathbb{E}}_{X,g} \mathrm{div}_g\left\{\mathop{\mathbb{E}}_{x \sim \mu}\left[\frac{1}{2\sqrt{\lambda}}x\right]\right\}$

$$= \frac{1}{2}\mathop{\mathbb{E}}_{X,g}\mathop{\mathbb{E}}_{x \sim \mu}\|x\|^2 - \frac{1}{2}\underbrace{\mathop{\mathbb{E}}_{X,g}\left\|\mathop{\mathbb{E}}_{x \sim \mu}[x]\right\|^2}_{=\mathbb{E}\|\mathbb{E}(X|Y)\|^2}$$

So

$$\frac{\partial}{\partial \lambda}F(\lambda) = \frac{1}{2}\mathbb{E}\|\mathbb{E}(X|Y)\|^2 \quad \text{as}$$

desired

Prop: For $\mu = \Pr[X = \cdot \mid Y]$,

$$\frac{\partial}{\partial g_i}\mathop{\mathbb{E}}_{x \sim \mu}[x_i] = \sqrt{\lambda}\left(\mathop{\mathbb{E}}_{x \sim \mu}[x_i^2] - \mathop{\mathbb{E}}_{x \sim \mu}[x_i]^2\right)$$

$$f'$$

Pf: Note

$$\mathbb{E}_{x\sim\mu}\left[x_i\right] = \frac{\int e^{-\frac{\lambda}{2}\|x\|^2 + \sqrt{\lambda}\langle x, y\rangle}\, x_i\, dx}{\int e^{-\frac{\lambda}{2}\|x\|^2 + \sqrt{\lambda}\langle x, y\rangle}\, dx}$$

$$= \frac{\int e^{-\frac{\lambda}{2}\|x\|^2 + \langle x, X\rangle + \sqrt{\lambda}\langle x, g\rangle}\, x_i\, dx}{\int e^{-\frac{\lambda}{2}\|x\|^2 + \langle x, X\rangle + \sqrt{\lambda}\langle x, g\rangle}\, dx},$$

So

$$\frac{\partial}{\partial g_i}\mathbb{E}_{x\sim\mu}\left[x_i\right] = \frac{\sqrt{\lambda}\int e^{-\frac{\lambda}{2}\|x\|^2 + \langle x, X\rangle + \sqrt{\lambda}\langle x, g\rangle}\, x_i^2\, dx}{\int e^{-\frac{\lambda}{2}\|x\|^2 + \langle x, X\rangle + \sqrt{\lambda}\langle x, g\rangle}\, dx}$$

$$- \frac{\sqrt{\lambda}\left(\int e^{-\frac{\lambda}{2}\|x\|^2 + \langle x, X\rangle + \sqrt{\lambda}\langle x, g\rangle}\, x_i\, dx\right)^2}{\left(\int e^{-\frac{\lambda}{2}\|x\|^2 + \langle x, X\rangle + \sqrt{\lambda}\langle x, g\rangle}\, dx\right)^2}$$

$$= \sqrt{\lambda}\left(\mathbb{E}_{\mu}\left[x_i^2\right] - \mathbb{E}_{\mu}\left[x_i\right]^2\right). \qquad \square$$