Hypothesis testing.

"Inherting" \( Y = \frac{1}{n} \sum x_i x_i^T + W \) \( P_x \)

"Dull" \( Y = W \) \( P_o \)

\( x \in \mathbb{R}^n \), \( x_i \sim P_0 \) zero mean, unit variance.

Interested in distinguishing \( P_x \) from \( P_o \).

General theory:

A test \( T : \mathbb{R}^{n \times n} \rightarrow \{ 0, 1 \} \).

Type I: \( P_o (T = 1) \)

Type II: \( P_x (T = 0) \).

\( \text{err}(T) = P_o (T = 1) + P_x (T = 0) \).

Interested in minimizing \( \text{err}(T) \).

Neyman-Pearson Lemma: the best minimizing \( T \rightarrow \text{err}(T) \) is

\[
T(y) = \begin{cases} 
1 & \text{if } L(y) > 1 \\
0 & \text{otherwise}
\end{cases}
\]

Likelihood ratio test.
\[ L(y) = \frac{dP_x}{dP_0}(y) : \text{likelihood ratio} \]

\[ \text{err}(T) = P_0[T = 0] + P_x[T = 1], \]

\[ \geq \int (\text{ALT} + L \cdot \text{ALT}) d\Omega. \]

\[ T^* = \text{ALT} \geq 1 \]

\[ \Rightarrow = 1 - d_{TV}(P, \Omega). \]

\[ d_{TV}(P, \Omega) = \frac{1}{2} \int |L - 1| d\Omega. \]

- Strong detection: \( \text{err}(T^*) \rightarrow 0 \)
  equivalently \( d_{TV}(P, \Omega) \rightarrow 1 \).

- Weak detection: \( \overline{\text{err}}(T^*) < 1 \).

- KL divergence:

\[ d_{KL}(P, \Omega) = E_P \left[ \log \frac{dP}{d\Omega} \right] \]

\[ = E_P \left[ \log L \right] \]

\[ = \mathbb{E}_\Omega \left[ L \log L \right] > 0. \]

\( \log x < x - 1 \)
\[ d_{KL}(P, Q) = \mathbb{E}_{Q} \left[ L \log L \right] \]
\[ \leq \mathbb{E}_{Q} \left[ L (L - 1) \right] \]
\[ = \mathbb{E}_{Q} \left[ L^2 \right] - 1 \]
\[ = \mathbb{E}_{Q} \left[ (L - 1)^2 \right] = d_{KL}(P, Q). \]

When strong detection not possible?

**Def (Contiguity):** \( P_n \rightarrow Q_n \) if for every sequence of events \((A_n)\) we have

\[ Q_n(A_n) \rightarrow 0 \Rightarrow P_n(A_n) \rightarrow 0. \]

**Mutual contiguity:** \( P_n \leq Q_n \) if \( P_n \rightarrow Q_n \) and \( Q_n \rightarrow P_n \).

**Fact:** If \( P_n \leq Q_n \) the strong detection is impossible

\[ A_n = \{ T = 1 \} \]

Suppose \( Q_n(A_n) \rightarrow 0 \Rightarrow P_n(A_n) \rightarrow 0 \).

\[ \Rightarrow \text{err}_n(T_1) \rightarrow 1. \]
• Proving contiguity: \( P_n(\mathcal{A}_n) \to 0 \).

\[
P_n(\mathcal{A}_n) = \mathbb{E}_{P_n} \left[ \sum_{i \leq n} I_{\mathcal{A}_n^i} \right].
\]

\[
= \mathbb{E}_{P_n} \left[ \sum_{i \leq n} I_{\mathcal{A}_n^i} \right].
\]

\[
\leq \mathbb{E}_{P_n} \left[ \sum_{i \leq n} I_{\mathcal{A}_n^i} \right].
\]

\[
\leq \mathbb{E}_{P_n} \left[ \sum_{i \leq n} I_{\mathcal{A}_n^i} \right].
\]

\[
= \mathbb{E}_{P_n} \left[ \sum_{i \leq n} I_{\mathcal{A}_n^i} \right].
\]

If \( \sup_n \mathbb{E}_{P_n} \left[ \sum_{i \leq n} I_{\mathcal{A}_n^i} \right] < \infty \), then

\[
P(\mathcal{A}_n) \to 0 \quad \implies \quad P_n \to P_n^1.
\]

Remark: \( d_{KL} (P_n, P_n^1) \leq \mathbb{E}_{P_n} \left[ \sum_{i \leq n} I_{\mathcal{A}_n^i} \right] - 1 \).

• Consider spiked wigener:

\[
L = \frac{dP_n}{dP_0}, \quad P_n = P_{\lambda}, \quad Q_n = P_0.
\]

\[
Y_{ij} = \frac{\lambda}{n} x_i x_j + \omega_{ij}, \quad i \neq j,
\]

\[
Y_{ii} = \frac{\lambda}{n} x_i^2 + \omega_{ii}, \quad i.
\]

\[
dP_\lambda(X) = \int \frac{1}{(2\pi)^{n/2}} e^{-\frac{1}{2} \sum_{ij} (Y_{ij} - \frac{\lambda}{n} x_i x_j)^2} dP_0^n(X), \quad \int dY.
\]

\[
L = \int e^{\sum_{i \neq j} \frac{\lambda}{n} x_i x_j} \left( \sum_{i \neq j} y_i y_j - \frac{1}{n} \sum_{i} y_i^2 \right) dP_0^n(X).
\]

(Ignoring the diagonal)
Including the diagonal: \( w_i \sim N(0, \sigma^2) \).

\[
L = \int e^{\frac{1}{\sigma^2} \sum_i y_i x_i x_i - \frac{1}{n} x_i^2} \left( 1 + \frac{1}{\sigma^2} \sum_i y_i x_i^2 - \frac{1}{2n} x_i^4 \right) dP_0^n(x)
\]

Posterior measure:

\[
\Pi(\theta | y) = \frac{e^{\eta \theta(x)}}{Z_n}
\]

\[\Rightarrow L = Z_n \]

\[
d_{KL}(P(x), P_0) = E_{P_0} \left[ \log Z_n(y) \right]
\]

\[= F_n(\lambda).\]

we know:

\[
\frac{1}{n} d_{KL}(P(x), P_0) \rightarrow \phi_{RS}(\lambda) \rightarrow 0
\]

\[
\frac{1}{n} E_{P_0} \log L
\]

\[
\begin{cases}
\phi_{RS}(\lambda) > 0 & \text{if } \lambda > \lambda c \\
\phi_{RS}(\lambda) = 0 & \text{if } \lambda \leq \lambda c
\end{cases}
\]

Second moment bounded \( \Rightarrow d_{KL}(P(x), P_0) \) bounded

\[\Rightarrow \phi_{RS}(\lambda) = 0\]

\[\Rightarrow \sup \frac{1}{n} E_{P_0} \left[ L(y)^2 \right] < \infty\]

then \( \lambda < \lambda c \).
Conversely, there are scenarios where $\lambda < \lambda_c$ but

$$\mathbb{E}_\lambda \left[ \mathbb{E}_\lambda \right] \to +\infty .$$

Conditioning the second moment computation:

Identify a sequence of good events $E_n$ such that $\mathbb{P}_n(E_n) \to 1$

I want to prove $\mathbb{P}_n < \infty$.

Take a sequence of events $(A_n) \text{ s.t. } \mathbb{P}_n(A_n) \to 0$

$$\mathbb{P}_n(A_n) = \mathbb{P}_n(A_n | E_n) \mathbb{P}(E_n) + \mathbb{P}_n(A_n | E_n^c) \mathbb{P}(E_n^c)$$

$$\leq \mathbb{P}_n(A_n | E_n) + \mathbb{P}(E_n^c) .$$

$$\leq \mathbb{E} \mathbb{P}_n \left[ \mathbb{E}_n \right] \frac{1}{2} \mathbb{P}_n(A_n) + \mathbb{E} \mathbb{P}_n(E_n^c) .$$

For the spiked wigner model:

- Can we prove strong detection for all $\lambda > \lambda_c$?

- Can we prove mutual contiguity for all $\lambda < \lambda_c$?

and what is the error of the best test for $\lambda \leq \lambda_c$?

Remark: unclear what happens at $\lambda = \lambda_c$ ...

The guess: detection is possible??

$\lambda \sim n^a$ ??
Strong detection above \( \lambda_c \): 

we have \( \frac{1}{n} E_{P_0} \log L \to \Phi_{P_0}(\lambda) \).

\[ \Phi_{P_0}(\lambda) > 0 \text{ if } \lambda > \lambda_c. \]

\[ E_{P_0} \log L \approx \log E_{P_0} [I] = 0. \]

\[ \lim \frac{1}{n} E_{P_0} \log L \equiv 0. \]

Look at \( \frac{1}{n} \log L \) and decide based on its sign.

Concentration of measure: (prior has bounded support)

\[ P_{\lambda} \left( \left| \frac{1}{n} \log L - E_{\lambda} \log L \right| > t \right) \leq e^{-\frac{nt^2}{2}} \quad \forall \lambda > 0. \]

If \( \lambda > \lambda_c : \frac{1}{n} \log L \equiv \Phi_{P_0}(\lambda) \) w.p. \( \to 1 \)

under \( P_\lambda \)

under \( P_0 : \frac{1}{n} \log L \leq \varepsilon \) w.p. \( \to 1 \)

\( \forall \varepsilon > 0 \).

The test \( T = \frac{1}{2} \left( \frac{1}{n} \log L \geq \frac{1}{2} \Phi_{P_0}(\lambda) \right) \) achieves strong detection.
Remark: \( \frac{1}{n} \mathbf{E}_0 \log L \rightarrow \Phi_{\text{Par}}(\lambda) \neq \lambda \geq 0 \)

\[ \sum \text{Talagrand} (\text{Rademacher prior}) \xrightarrow{\text{a.s.}} \mathcal{F} \]

\[ \phi_{\mathcal{F}}(\lambda) \]

\[ \lambda_c \leq \overline{\lambda}_e \]

**Question:** Is \( \lambda_c = \overline{\lambda}_e \) for all priors?

Yes if \( \mathbf{P}_0 = \text{Rad} (\frac{1}{2}) \), \( \lambda_c = \overline{\lambda}_e = 1 \).

(Example?)

**Difficult to compute in general.**

What about efficient tests?

Spectral tests?

*First eigenvalue of \( Y \)?*

Under \( \mathbf{P}_0: Y = \mathbf{w} \).

\[ \lambda_1(Y) \rightarrow 2 \text{ a.s.} \]

Under \( \mathbf{P}_A: Y = \mathbf{A} \mathbf{x}^T + \mathbf{w} \).
Fetal - Peche 2007:
\[ \lambda_1(Y) \rightarrow \begin{cases} 2 & \text{if } \lambda < 1 \\ \sum \lambda + \frac{1}{\lambda} & \text{otherwise} \end{cases} \]

The first eigenvalue achieves strong detection for \( \lambda > 1 \).

What about \( \lambda \in (1, 1) \)?

Proposition:

- In fact we can show that any test based on the eigenvalues of \( Y \) fails at strong detection for \( \lambda < 1 \).

\[ Y = \sqrt{\frac{1}{n}} X \tilde{X}^T + W \]

Let \( R \) be orthogonal matrix such which sends \( X \) to the all ones vector

\[ Q X = p_n \tilde{X} \quad p_n = \frac{1 \times n^2}{n} \rightarrow 1 \]

\[ \tilde{Y} = R \tilde{X} R^T \]

\[ = \sqrt{\frac{1}{n}} p^2_n \quad \tilde{X} \tilde{X}^T + Q W Q^T \]

\[ = 1 + o(1) \]

The eigenvalues of \( \tilde{Y} = \) The eig. vals (\( Y \)).
\( \lambda_c = 1 \) for \( \bar{Y} \).

So law of \( \bar{Y} \) is contiguously to law of \( \bar{W} \).

\( \Rightarrow \) no test operating on the eigenvalues succeeds for \( \lambda < 1 \).

we will prove that \( \bar{P}_x = \bar{P}_0 \) for \( \lambda < 1 \).

for Rademacher.

Easier to prove contiguity up to \( \lambda = \frac{1}{\tilde{\alpha}^2} \).

where \( \tilde{\alpha}^2 \) is subgaussian parameter of the prior.

for Rademacher: \( \tilde{\alpha}^2 = 1 \). [via the second mentioned \( \bar{J} \).]