Marchenko–Pastur Distribution

Given $X \in \mathbb{R}^{p*n}, X_i \sim N(0, I_p) i.i.d.$ Calculate sample covariance: $\hat{\Sigma} = \sum_{k=1}^{p} \hat{l_k} \hat{u_k} \hat{u_k}^T$ where $\hat{l}_k$ are eigenvalues of covariance matrix $\hat{\Sigma}$ Consider the random measure: $\mu_p(A) = \frac{1}{p} \#\{\hat{l}_k \in A\}$

Theorem. (Marchenko–Pastur Distribution) *Assume $p, n \to \infty$ so that the ratio $p/n \to \gamma$. Then, $\mu_p \to d\mu$ almost everywhere, where*

$$\mu(A) = \begin{cases} (1 - \frac{1}{\lambda} \mathbf{1}_{0 \in A} + \nu(A), if \gamma > 1 \\ \nu(x), if \gamma \leq 1 \end{cases}$$

, and $d\nu(A) = \frac{1}{2\pi\sigma^2} \frac{\sqrt{(b-x)(x-a)}}{\gamma x} \mathbf{1}_{[]}$ *, where*

$$\begin{cases} a = \sigma^2 (1 - \sqrt{\gamma})^2 \\ b = \sigma^2 (1 + \sqrt{\gamma})^2 \end{cases}$$

Remark (Marchenko–Pastur Distribution)**.** *Marchenko–Pastur Distribution gives an asymptotic distribution function to estimate the distribution of eigenvalue of a normal random matrix in the case that the ratio $p/n \to \gamma$ doesn't converge to zero. We can observe the measure function $d\nu(x)$ and find that the largest eigenvalue of this random matrix with normal elements will be larger than σ^2. In the following example, we will show how this phenomenon affect original data and how can we use it to identify real eigenvalue of data.*

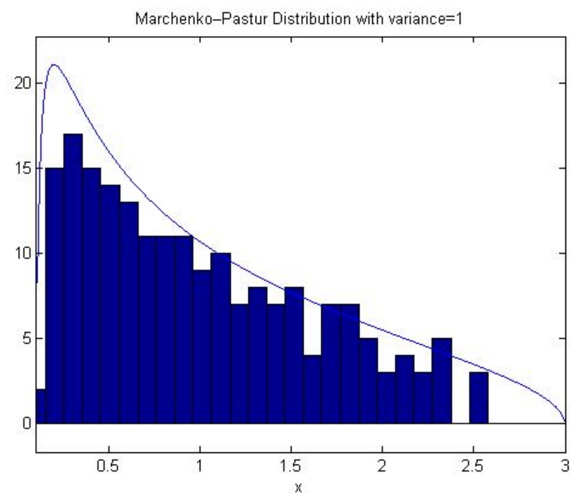

Figure 1: A simple Marchenko–Pastur Distribution with $\sigma^2 = 1$

Marchenko–Pastur Distribution with Rank one Signal

In the last paragraph, we briefly introduced the Marchenko–Pastur Distribution. Now let's see how Marchenko–Pastur Distribution can be useful in real signal analysis.

Now imagine a n inputs signal X in dimension p with only one rank, which means there's only one entry of X is non-zero. And the variance of the value of that entry is σ_x^2. However, on the way of transmission, noises will occur in every entry of signal X. As a result, the signal Y the observer received is different from the original signal X. For our analysis convenient, here we suppose the noises in each entry of signal is $i.i.d$ and their densities are normal distribution with variance σ^2. So, we have:

$$\begin{aligned} &X \in \mathbb{R}^{n\times p} \\ &Y = X + \xi, \qquad \xi \sim N(0, \sigma^2 I_p) \\ &\frac{p}{n} \to \gamma > 0, \qquad \gamma < 1 \end{aligned}$$

Since X is rank 1, we can decompose X as the following,

$$X = \sigma_x^2 u u^T$$

So we can decompose Y as follow,

$$Y = X + \xi \sim N(0, \sigma_x^2 u u^T + I_p)$$

Look at the ith component of Y,

$$\begin{aligned} &Y_i = \Sigma^{1/2} Z_i \\ &Z_i = N(0, I_p) \\ &\Sigma = U \begin{bmatrix} 1+\sigma_x^2 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix} U^T \end{aligned}$$

Covariance matrix,

$$\begin{aligned} \hat{\Sigma_n} &= \frac{1}{n-1} Y Y^T \\ &= \Sigma^{1/2} \frac{1}{n-1} Z Z^T \Sigma^{1/2} \\ &= \Sigma^{1/2} S_n \Sigma^{1/2} \end{aligned}$$

Where S_n is the sample covariance matrix

$$S_n = \frac{1}{n-1} Z Z^T$$

Then look at the spectral decomposition of S_n

$$S_n = W \Lambda W^T, \qquad \Lambda = diag(\lambda_1, \ldots, \lambda_p)$$

Consider the covariance estimator, we can find out that it is similar to $S_n \Sigma$,

$$\hat{\Sigma_n} = \Sigma^{1/2} S_n \Sigma^{1/2} \sim S_n \Sigma$$

Also, we know that when two matrices are similar, they share the same eigenvalues,

$$\sigma(\hat{\Sigma_n}) = \sigma(S_n \Sigma)$$

Take the top eigenpair of of $S_n \Sigma$ from $\hat{\Sigma_n}$

$$(\lambda, v)$$

And we assume that λ escape the Marchenko–Pastur Distribution of the noises

$$\lambda \notin \sigma(S_n)$$

Following is the eigenvalue histogram of a one sparse signal whose variance $\sigma_x = 2$ with noises' variance $\sigma = 1$

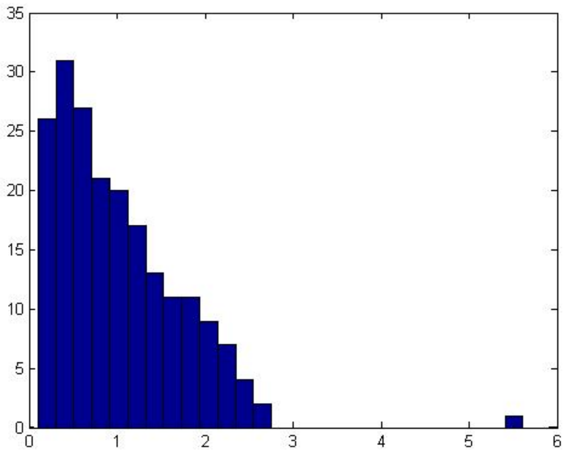

Figure 2: Rank 1 signal whose variance $\sigma_x = 2$ with noises' variance $\sigma = 1$

As a result, we know that this eigenpair must have great relation with the original signal. So the next problem to solve is that how can we extract the information from the top eigenvector to recover original signal?