Detection of Signal in the Spiked Rectangular Models

Ji Hyung Jung, Hye Won Chung, and Ji Oon Lee

Korea Advanced Institute of Science and Technology (KAIST)

2021 ICML

Spiked rectangular models

- Rectangular matrix with *spiked mean* (additive model)

$$\sqrt{\lambda}\boldsymbol{u}\boldsymbol{v}^T + X$$

X: $M \times N$ random i.i.d. matrix, centered with variance N^{-1}

$\boldsymbol{u} \in \mathbb{R}^M, \boldsymbol{v} \in \mathbb{R}^M, \|\boldsymbol{u}\| = \|\boldsymbol{v}\| = 1$

- Rectangular matrix with *spiked covariance* (multiplicative model)

$$(I + \lambda\boldsymbol{u}\boldsymbol{u}^T)^{1/2}X$$

X: $M \times N$ random i.i.d. matrix, centered with variance N^{-1}

$\boldsymbol{u} \in \mathbb{R}^M, \|\boldsymbol{u}\| = \|\boldsymbol{v}\| = 1$

Spiked rectangular models

$$\sqrt{\lambda}\boldsymbol{u}\boldsymbol{v}^T + X \qquad (I + \lambda \boldsymbol{u}\boldsymbol{u}^T)^{1/2} X$$

- If $\boldsymbol{u}$ and $\boldsymbol{v}$ are centered, the population covariance $\Sigma = I + \lambda \boldsymbol{u}\boldsymbol{u}^T$.
- If $\boldsymbol{v}$ is i.i.d. Gaussian (scaled standard normal), the two models coincide.

- High-dimensional assumption $M, N \to \infty$ with $M/N \to d_0 \in (0, \infty)$.
- If $\lambda > \sqrt{d_0}$, signal can be reliably detected and recovered by PCA. (BBP transition)
- If $0 < \lambda < \sqrt{d_0}$ and the noise is Gaussian,
  - signal cannot be detected by PCA,
  - no tests can reliably detect the signal.
    (Onatski, Moreira, Hallin '13, '14, El Alaoui, Jordan '18)

Improved PCA - additive model

$Y = \sqrt{\lambda}\boldsymbol{u}\boldsymbol{v}^T + X$, $\sqrt{N}X_{ij}$ is drawn from a distribution with a density g.

- Fisher information

$$F_g = \int_{-\infty}^{\infty} \frac{g'(x)^2}{g(x)} \mathrm{d}x.$$

($F_g \geq 1$ with equality if and only if g is Gaussian.)

- Entrywise transformation

$$h(x) := -\frac{g'(x)}{g(x)}, \qquad \widetilde{Y}_{ij} = \frac{1}{\sqrt{F_g N}} h(\sqrt{N} Y_{ij}).$$

Theorem (J.-C.-L.)

For the largest eigenvalue $\widetilde{\mu}_1$ of $\widetilde{Y}\widetilde{Y}^T$, almost surely

$$\widetilde{\mu}_1 \to \begin{cases} (1+\lambda_g)(1+\frac{d_0}{\lambda_g}) & \text{if } \lambda_g > \sqrt{d_0}, \\ (1+\sqrt{d_0})^2 & \text{if } \lambda_g < \sqrt{d_0}. \end{cases} \qquad (\lambda_g = \lambda F_g)$$

Improved PCA - multiplicative model

$Y = (I + \lambda \boldsymbol{u}\boldsymbol{u}^T)^{1/2} X$, $\lambda = 2\gamma + \gamma^2$.

- Entrywise transformation

$$h_\alpha(x) := -\frac{g'(x)}{g(x)} + \alpha x, \quad \alpha = \frac{-\gamma F_g + \sqrt{4F_g + 4\gamma F_g + \gamma^2 F_g^2}}{2(1+\gamma)}.$$

$$\widetilde{Y}_{ij} \equiv \widetilde{Y}_{ij}^{(\alpha)} = \frac{1}{\sqrt{(\alpha^2 + 2\alpha + F_g)N}} h_\alpha(\sqrt{N} Y_{ij})$$

Theorem (J.-C.-L.)

For the largest eigenvalue $\widetilde{\mu}_1$ of $\widetilde{Y}\widetilde{Y}^T$, almost surely

$$\widetilde{\mu}_1 \to \begin{cases} (1+\lambda_g)(1+\frac{d_0}{\lambda_g}) & \text{if } \lambda_g > \sqrt{d_0}, \\ (1+\sqrt{d_0})^2 & \text{if } \lambda_g < \sqrt{d_0}. \end{cases} \quad (\lambda_g = \gamma + \gamma^2 F_g + \gamma(1+\gamma)\alpha \geq \lambda)$$

Reconstruction by the proposed PCA

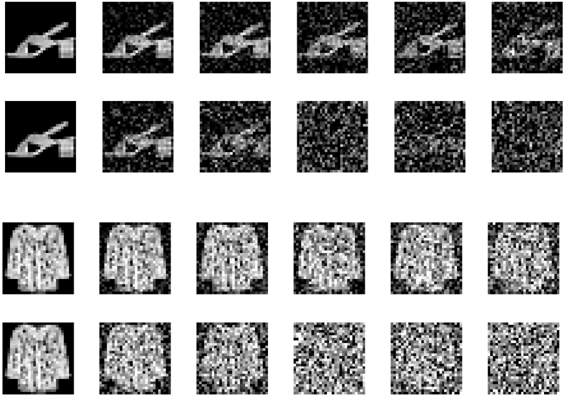

Figure: Reconstruction performance of the proposed PCA (top lines) and the standard PCA (bottom lines) for two FashionMNIST images, with $N = [3136, 1568, 784, 588, 392]$ and $M = 784$.

Weak detection of signal

Recall $\mathbb{E}[X_{ij}] = 0$, $\mathbb{E}[X_{ij}^2] = 1/N$. Set $\mathbb{E}[X_{ij}^4] = w_4/N^2$.

- Hypothesis testing $\quad \boldsymbol{H}_0 : \lambda = 0, \quad \boldsymbol{H}_1 : \lambda = \omega > 0$
- Compute the test statistic

$$L_\omega = -\log\det\left(\left(1+\frac{d_0}{\omega}\right)(1+\omega)I - YY^T\right) + \frac{\omega}{d_0}\left(\frac{2}{w_4-1}-1\right)(\operatorname{Tr} YY^T - M) + M\left[\frac{\omega}{d_0} - \log\left(\frac{\omega}{d_0}\right) - \frac{1-d_0}{d_0}\log(1+\omega)\right].$$

Theorem (J.-C.-L.)

For both the additive model and the multiplicative model,

$$L_\omega \Rightarrow \mathcal{N}(m(\lambda), V_0)$$

Hypothesis testing - Algorithm

- Set $m_\omega = \frac{1}{2}(m(0) + m(\omega))$.
- Accept $\boldsymbol{H}_0$ if $L_\omega \le m_\omega$. Accept $\boldsymbol{H}_1$ if $L_\omega > m_\omega$.
- The error of the test

$$\mathrm{err}(\omega) = \mathbb{P}(L_\omega > m_\omega | \boldsymbol{H}_0) + \mathbb{P}(L_\omega \le m_\omega | \boldsymbol{H}_1) \to \mathrm{erfc}\left(\frac{\sqrt{V_0}}{4\sqrt{2}}\right)$$

- Universality: For any deterministic or random $\boldsymbol{u}$ (and $\boldsymbol{v}$), the proposed test and its error do not change,
- Optimality: For Gaussian noise, the error of the proposed test with low computational complexity converges to the optimal limit.

Error of the proposed test

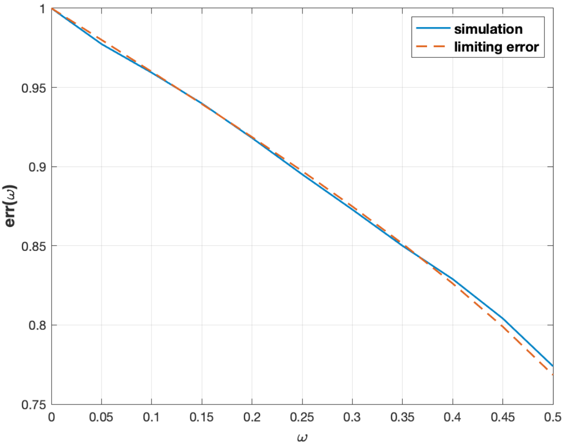

Figure: The error from the simulation (solid) and the theoretical error (dashed).

Conclusion / Future Works

- We showed that PCA can be improved for non-Gaussian noise by transforming the data entrywise.
- We proved the effective SNR and the optimal entrywise transforms for both the additive model and the multiplicative model.
- We proposed a universal hypothesis test for the weak detection, which is optimal if the noise is Gaussian.
- Future work: improving the hypothesis test by applying the entrywise transformation