Applications of Random Matrix Theory in Wireless Underwater Communication

Why Signal Processing and Wireless Communication Need Random Matrix Theory

Atulya Yellepeddi

May 13, 2013

18.338- Eigenvalues of Random Matrices, Spring 2013 -Final Project

Signal Processing and the Law of Large Numbers

- Want to estimate a "population size" m from n observations
  - Methods and Analysis based on statistics of population as $n \to \infty$
  - Usually works if $n \gg m$

Signal Processing and the Law of Large Numbers

- Want to estimate a "population size" m from n observations
  - Methods and Analysis based on statistics of population as $n \to \infty$
  - Usually works if $n \gg m$
- Modern signal processing- cases when $n \sim m$
  - Parameter of a time-varying system (eg., underwater communication)
  - Population size grows as observations grow (eg., social networks)
  - Huge population size (eg., huge array beamformers)

Signal Processing and the Law of Large Numbers

- Want to estimate a "population size" m from n observations
  - Methods and Analysis based on statistics of population as $n \to \infty$
  - Usually works if $n \gg m$
- Modern signal processing- cases when $n \sim m$
  - Parameter of a time-varying system (eg., underwater communication)
  - Population size grows as observations grow (eg., social networks)
  - Huge population size (eg., huge array beamformers)
- **Random Matrix Theory**: excellent at making predictions in such scenarios

Why Random Matrix Theory?

Why Random Matrix Theory?

For Mathematicians, $30 \approx \infty$

So with reasonable population sizes $m \sim 30$ or more, we can make highly accurate predictions which work for small number of observations.

Why Random Matrix Theory?

For Mathematicians, $30 \approx \infty$

So with reasonable population sizes $m \sim 30$ or more, we can make highly accurate predictions which work for small number of observations.

Almost Anything is IID Gaussian

Results for IID Gaussian ensembles carry over, in practice, to all sorts of ensembles (with some caveats!) if they are "reasonably" like Gaussian, and "more or less" independent.

Least Squares Channel Estimation- the Problem

- Transmit $\boldsymbol{u}(n)\dots$
  - $m \times 1$ vectors, independent from time to time
  - $\mathbb{E}[\boldsymbol{u}(n)] = 0$
  - $\mathbb{E}[\boldsymbol{u}(n)\boldsymbol{u}^{\dagger}(m)] = \boldsymbol{R}\delta(n-m)$

Least Squares Channel Estimation- the Problem

- Transmit $\boldsymbol{u}(n)$...
  - $m \times 1$ vectors, independent from time to time
  - $\mathbb{E}[\boldsymbol{u}(n)] = 0$
  - $\mathbb{E}[\boldsymbol{u}(n)\boldsymbol{u}^{\dagger}(m)] = \boldsymbol{R}\delta(n-m)$
- ...across channel $\boldsymbol{w}_0$...
  - outputs: $d(n) = \boldsymbol{w}_0^{\dagger}\boldsymbol{u}(n) + v(n)$
  - $v(n)$ is noise of power σ_v^2

Least Squares Channel Estimation- the Problem

- Transmit $\boldsymbol{u}(n)$...
  - $m \times 1$ vectors, independent from time to time
  - $\mathbb{E}[\boldsymbol{u}(n)] = 0$
  - $\mathbb{E}[\boldsymbol{u}(n)\boldsymbol{u}^\dagger(m)] = \boldsymbol{R}\delta(n-m)$
- ...across channel $\boldsymbol{w}_0$...
  - outputs: $d(n) = \boldsymbol{w}_0^\dagger \boldsymbol{u}(n) + v(n)$
  - $v(n)$ is noise of power σ_v^2
- ...what's $\boldsymbol{w}_0$?
  - Don't know $\boldsymbol{R}$

Least Squares Channel Estimation- the Problem

- Transmit $\boldsymbol{u}(n)$...
  - $m \times 1$ vectors, independent from time to time
  - $\mathbb{E}[\boldsymbol{u}(n)] = 0$
  - $\mathbb{E}[\boldsymbol{u}(n)\boldsymbol{u}^\dagger(m)] = \boldsymbol{R}\delta(n-m)$
- ...across channel $\boldsymbol{w}_0$...
  - outputs: $d(n) = \boldsymbol{w}_0^\dagger \boldsymbol{u}(n) + v(n)$
  - $v(n)$ is noise of power σ_v^2
- ...what's $\boldsymbol{w}_0$?
  - Don't know $\boldsymbol{R}$

Least Squares Solution

$$\hat{\boldsymbol{w}}(n) = \underbrace{\boldsymbol{R}^{-1}(n)}_{=\sum_{i=1}^{n} \boldsymbol{u}(i)\boldsymbol{u}^\dagger(i)+\delta \boldsymbol{I}} \overbrace{\boldsymbol{z}(n)}^{=\sum_{i=1}^{n} u(i)d^*(i)} \tag{1}$$

Least Squares Channel Estimation-the Challenge

- Idea is as $n \to \infty$, the Sample Covariance Matrix $\frac{\boldsymbol{R}(n)}{n} \to \boldsymbol{R}$.

Least Squares Channel Estimation-the Challenge

- Idea is as $n \to \infty$, the Sample Covariance Matrix $\frac{\boldsymbol{R}(n)}{n} \to \boldsymbol{R}$.
- Performance: the channel estimation error $\epsilon(n) = \boldsymbol{w}_0 - \hat{\boldsymbol{w}}(n)$

Least Squares Channel Estimation-the Challenge

- Idea is as $n \to \infty$, the Sample Covariance Matrix $\frac{\boldsymbol{R}(n)}{n} \to \boldsymbol{R}$.
- Performance: the channel estimation error $\epsilon(n) = \boldsymbol{w}_0 - \hat{\boldsymbol{w}}(n)$
- If $n \to \infty$ (practically, $n \gg m$) is assumed:

$$\mathbb{E}[\|\epsilon(n)\|_2^2] \approx \frac{1}{n}\sigma_v^2 tr\{\mathbf{R}^{-1}\} \tag{2}$$

Least Squares Channel Estimation-the Challenge

- Idea is as $n \to \infty$, the Sample Covariance Matrix $\frac{\boldsymbol{R}(n)}{n} \to \boldsymbol{R}$.
- Performance: the channel estimation error $\epsilon(n) = \boldsymbol{w}_0 - \hat{\boldsymbol{w}}(n)$
- If $n \to \infty$ (practically, $n \gg m$) is assumed:

$$\mathbb{E}[||\epsilon(n)||_2^2] \approx \frac{1}{n}\sigma_v^2 tr\{\mathbf{R}^{-1}\} \tag{2}$$

- **Bad predictions for $n \sim m$**
  - Occurs in applications like underwater acoustic communication, where channel varies quickly
  - Or when taking observations is expensive

Least Squares Channel Estimation-the Challenge

- Idea is as $n \to \infty$, the Sample Covariance Matrix $\frac{\boldsymbol{R}(n)}{n} \to \boldsymbol{R}$.
- Performance: the channel estimation error $\epsilon(n) = \boldsymbol{w}_0 - \hat{\boldsymbol{w}}(n)$
- If $n \to \infty$ (practically, $n \gg m$) is assumed:

$$\mathbb{E}[\|\epsilon(n)\|_2^2] \approx \frac{1}{n}\sigma_v^2 tr\{\mathbf{R}^{-1}\} \tag{2}$$

- **Bad predictions for $n \sim m$**
  - Occurs in applications like underwater acoustic communication, where channel varies quickly
  - Or when taking observations is expensive
- Random Matrix Theory allows better predictions

The Random Matrix Theory Results

Let $M_k(m,n) = \frac{1}{m}\mathbb{E}\left[\mathrm{Tr}\left(\boldsymbol{R}^{-k}(n)\right)\right]$. Then...

The Random Matrix Theory Results

Let $M_k(m,n) = \frac{1}{m}\mathbb{E}\left[\mathrm{Tr}\left(\boldsymbol{R}^{-k}(n)\right)\right]$. Then...

Channel Estimation Error:

$$\mathbb{E}\left[\|\epsilon(n)\|_2^2\right] = m\left(\sigma_v^2 M_1(m,n) + \delta^2 M_2(m,n) - \delta\sigma_v^2 M_2(m,n)\right) \quad (3)$$

Moments: Independent Input Observations

Let $\boldsymbol{A} = \frac{1}{n}\boldsymbol{X}\boldsymbol{X}^{\dagger}\boldsymbol{T}^{1/2}$.
2 key ideas to compute the moments:

Moments: Independent Input Observations

Let $\boldsymbol{A} = \frac{1}{n}\boldsymbol{X}\boldsymbol{X}^{\dagger}\boldsymbol{T}^{1/2}$.

2 key ideas to compute the moments:

- Assume m, n are large but $c = \frac{m}{n}$ is a constant. Then, eigenvalue density of $\boldsymbol{R}(n)$ can be approximated (Marcenko Pastur law):

$$\mu_{\boldsymbol{R}(n)}(x) \approx \mu_{\boldsymbol{\Phi}}(x) = \frac{1}{n}\mu_{\boldsymbol{A}}\left(\frac{x-\delta}{n}\right) \tag{4}$$

Moments: Independent Input Observations

Let $\boldsymbol{A} = \frac{1}{n}\boldsymbol{X}\boldsymbol{X}^{\dagger}\boldsymbol{T}^{1/2}$.

2 key ideas to compute the moments:

- Assume m, n are large but $c = \frac{m}{n}$ is a constant. Then, eigenvalue density of $\boldsymbol{R}(n)$ can be approximated (Marcenko Pastur law):

$$\mu_{\boldsymbol{R}(n)}(x) \approx \mu_{\boldsymbol{\Phi}}(x) = \frac{1}{n}\mu_{\boldsymbol{A}}\left(\frac{x-\delta}{n}\right) \tag{4}$$

- Use the following to compute moments:

$$\lim_{m\to\infty}\frac{1}{m}\mathbb{E}[\mathsf{Tr}\left(\boldsymbol{B}^{k}\right)] = \int t^{k}\mu_{\boldsymbol{B}}(t)\,\mathrm{d}t \tag{5}$$

Moments: Independent Input Observations

Let $\boldsymbol{A} = \frac{1}{n}\boldsymbol{X}\boldsymbol{X}^{\dagger}\boldsymbol{T}^{1/2}$.

2 key ideas to compute the moments:

- Assume m, n are large but $c = \frac{m}{n}$ is a constant. Then, eigenvalue density of $\boldsymbol{R}(n)$ can be approximated (Marcenko Pastur law):

$$\mu_{\boldsymbol{R}(n)}(x) \approx \mu_{\boldsymbol{\Phi}}(x) = \frac{1}{n}\mu_{\boldsymbol{A}}\left(\frac{x-\delta}{n}\right) \tag{4}$$

- Use the following to compute moments:

$$\lim_{m\to\infty}\frac{1}{m}\mathbb{E}[\mathrm{Tr}\left(\boldsymbol{B}^k\right)] = \int t^k \mu_{\boldsymbol{B}}(t)\,\mathrm{d}t \tag{5}$$

Compute Moments Using:

$$M_k(m,n) \approx \int t^{-k}\mu_{\boldsymbol{\Phi}}(t)\,\mathrm{d}t \tag{6}$$

Predictions Made- Gaussian Input

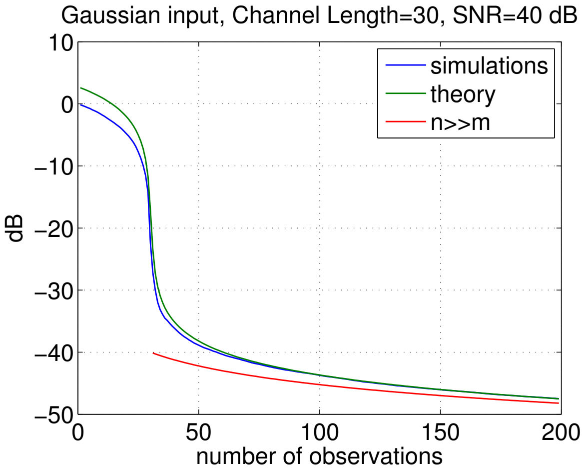

Figure : Channel Estimation MSE vs Number of Observations

Predictions Made- Gaussian Input

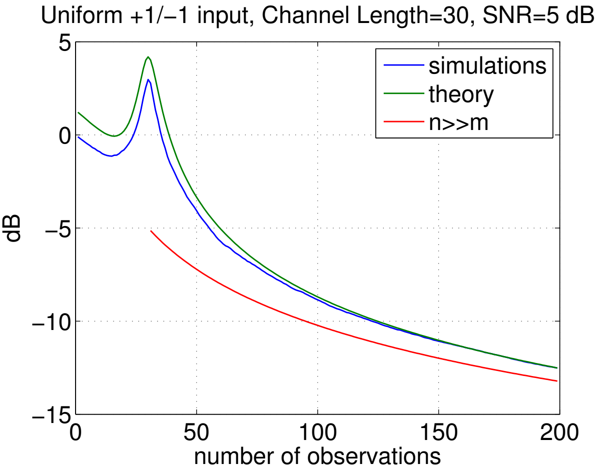

Figure : Channel Estimation MSE vs Number of Observations

Conclusions

- RMT makes nice predictions about signal processing systems running with a small number of observations
- Leads to identifying phenomena that were previously unknown
- Simple tools, but widely applicable
- More sophisticated tools available. . . how to use?