Non-Parametric Signal Detection and RMT

Boaz Nadler

Department of Computer Science and Applied Mathematics
The Weizmann Institute of Science

Joint work with Shira Kritchman
and with Federico Penna, Roberto Garello.

Oct. 2010

Talk Outline

1. Signal Detection Problems
2. Connection to Random Matrix Theory
3. Implications

Detection of signals embedded in noise

Given a measurement system with p sensors (antennas / microphones / hyperspectral camera / etc)

Observe multivariate samples $\mathbf{x}(t_j) \in \mathbb{R}^p$ of the form

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \sigma\xi(t) \qquad (*)$$

Detection of signals embedded in noise

Given a measurement system with p sensors (antennas / microphones / hyperspectral camera / etc)

Observe multivariate samples $\mathbf{x}(t_j) \in \mathbb{R}^p$ of the form

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \sigma\xi(t) \qquad (*)$$

where

- $\mathbf{s}(t) = (s_1(t), \ldots, s_K(t))'$ are K time-dependent signals.
- $\mathbf{A}$ is a $p \times K$ *fixed* unknown mixing matrix of rank K (steering matrix).

Detection of signals embedded in noise

Given a measurement system with p sensors (antennas / microphones / hyperspectral camera / etc)

Observe multivariate samples $\mathbf{x}(t_j) \in \mathbb{R}^p$ of the form

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \sigma\xi(t) \qquad (*)$$

where

- $\mathbf{s}(t) = (s_1(t), \ldots, s_K(t))'$ are K time-dependent signals.
- $\mathbf{A}$ is a $p \times K$ *fixed* unknown mixing matrix of rank K (steering matrix).
- σ - noise level, $\xi \sim \mathcal{N}(0, \mathbf{I}_p)$.

Problem Formulation

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \sigma\xi(t) \qquad (*)$$

Assume $\mathbf{s}(t) \in \mathbb{R}^K$ is stationary random process with a full rank covariance matrix, and that mixing matrix $\mathbf{A}$ is of rank K (e.g., there are indeed K identifiable sources).

Problem Formulation

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \sigma\xi(t) \qquad (*)$$

Assume $\mathbf{s}(t) \in \mathbb{R}^K$ is stationary random process with a full rank covariance matrix, and that mixing matrix $\mathbf{A}$ is of rank K (e.g., there are indeed K identifiable sources).

Problem Formulation

Given n i.i.d. samples $\mathbf{x}_i$ from the model (*), estimate the number of sources K.

Problem Setup

The linear mixture (factor) model

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \sigma\xi(t) \qquad (*)$$

appears in many different scientific fields.

Problem Setup

The linear mixture (factor) model

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \sigma\xi(t) \qquad (*)$$

appears in many different scientific fields.

- Analytical Chemistry / Chemometrics: **x** is the measured (logarithm of) spectra at p wavelengths, **s** - vector of concentrations of K chemical components. Eq. $(*)$ follows from Beer-Lambert's law.
- Signal Processing: **s** is a vector of K emitting sources, **x** - measurement at an array of p receivers (microphones, antennas, etc).

Problem Setup

The linear mixture (factor) model

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \sigma\xi(t) \qquad (*)$$

appears in many different scientific fields.

- Analytical Chemistry / Chemometrics: $\mathbf{x}$ is the measured (logarithm of) spectra at p wavelengths, $\mathbf{s}$ - vector of concentrations of K chemical components. Eq. $(*)$ follows from Beer-Lambert's law.
- Signal Processing: $\mathbf{s}$ is a vector of K emitting sources, $\mathbf{x}$ - measurement at an array of p receivers (microphones, antennas, etc).
- Statistical Modeling at large : linear mixture models / factor models / error in variables / two-way table with multiplicative interactions.

Rank Estimation - What is it needed for ?

In signal processing -

a preliminary step before blind source separation / independent component analysis, direction of arrival estimation, many other parametric procedures whose number of parameters depends on K - the number of sources.

Rank Estimation - What is it needed for ?

In signal processing -

a preliminary step before blind source separation / independent component analysis, direction of arrival estimation, many other parametric procedures whose number of parameters depends on K - the number of sources.

In chemometrics -

process control, optimal number of latent variables in regression/calibration models, first step prior to self modeling curve resolution and many other estimation procedures.

Nonparametric Detection

In this talk - focus on **nonparametric** detection.

Nonparametric Detection

In this talk - focus on **nonparametric** detection.

- No assumption on structure / smoothness / sparsity of the mixing matrix **A**. (no assumptions on array manifold structure)

Nonparametric Detection

In this talk - focus on **nonparametric** detection.

- No assumption on structure / smoothness / sparsity of the mixing matrix $\mathbf{A}$. (no assumptions on array manifold structure)

- No assumption on possible non-Gaussian / bi-modal / finite alphabet distribution of the random variables $\mathbf{s}(t)$.

Nonparametric Detection

In this talk - focus on **nonparametric** detection.

- No assumption on structure / smoothness / sparsity of the mixing matrix $\mathbf{A}$. (no assumptions on array manifold structure)

- No assumption on possible non-Gaussian / bi-modal / finite alphabet distribution of the random variables $\mathbf{s}(t)$.

In this setting, assuming Gaussian signals, eigenvalues $\ell_1 \geq \ell_2 \geq \ldots \geq \ell_p$ of the sample covariance matrix

$$\mathbf{S}_n = \sum_i \mathbf{x}_i' \mathbf{x}_i$$

are *sufficient statistics* for eigenvalues of the population covariance matrix $\mathbf{\Sigma}$ [James 66', Muirhead 78'].

Detection of Signals as a model selection problem

Given K sources, the population covariance matrix of the observations has diagonal form

$$\mathbf{W}'\mathbf{\Sigma}\mathbf{W} = \sigma^2\mathbf{I}_p + diag(\lambda_1, \lambda_2, \ldots, \lambda_K, 0, \ldots, 0) \qquad (1)$$

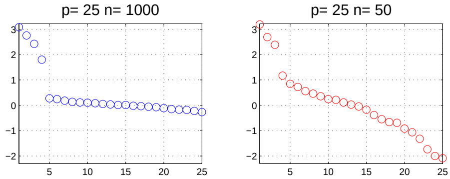
Model Selection Problem: Given $\{\ell_j\}_{j=1}^{p}$, determine which model of the form (1) is most likely.

Detection of Signals as a model selection problem

Given K sources, the population covariance matrix of the observations has diagonal form

$$\mathbf{W}'\boldsymbol{\Sigma}\mathbf{W} = \sigma^2\mathbf{I}_p + diag(\lambda_1, \lambda_2, \ldots, \lambda_K, 0, \ldots, 0) \qquad (1)$$

Model Selection Problem: Given $\{\ell_j\}_{j=1}^{p}$, determine which model of the form (1) is most likely.



Eigenvalue Based Rank Estimation

Challenge: Distinguish between large yet insignificant noise eigenvalues and small yet significant signal eigenvalues.

Problem Parameters: n - number of samples, p - dimensionality, σ - noise level.

In classical array processing $p/n \ll 1$,

In chemometrics and in some modern radar/sonar systems: $p/n = O(1)$ and often $p/n \gg 1$.

Previous Approaches

Can be divided into 2 main disciplines,

- Nested Hypothesis Tests (with various test statistics)
- Information Theoretic Criteria (BIC, MDL, AIC, etc)

Previous Approaches

The "godfathers"

- Bartlett (1940's), Lawley (1950's) - likelihood ratio tests, tests for sphericity, assume Gaussian observations, asymptotic expansions for p fixed, $n \to \infty$.

$$T = \frac{\left(\prod \ell_i\right)^{1/p}}{\frac{1}{p}\sum \ell_i}$$

This statistic does not work well when p is of the same order of n and is undefined if $p > n$.

Previous Approaches

In chemometrics ($p \gg n$):

- Malinowski F-test, 1977, 1980, 1987-1990, *Analytical Chemistry, J. Chemometrics*.
- **Faber and Kowalski, Modification of Malinowski's F-test, J. Chemometrics, 1997.**
- Faber, Buydens and Kateman, Aspects of pseudorank estimation methods based on the eigenvalues of principal component analysis of random matrices, 1994.
- at least 15 other papers describing and comparing various algorithms.

Previous Approaches

In signal processing and statistics literature:

- Wax & Kailath, *Detection of signals by information theoretic criteria*, 85'
- Zhao, Krishnaiah and Bai. JMVA, 1986
- Fishler, Grosmann and Messer, IEEE Sig. Proc. 2002.
- P-J. Chung, J.F. Böhme, C.F. Mecklenbraüker and A.O. Hero, 2007.
- many other papers in signal processing and in statistics.
- Silverstein & Combettes, 1992
- Schott, *A high-dimensional test for the equality of the smallest eigenvalues of a covariance matrix*, JMVA, 2006.
- Rao & Edelman, *Sample eigenvalue based detection of high-dimensional signals in white noise using relatively few samples*. 2007.

Detection of Signals by Information Theoretic Criteria

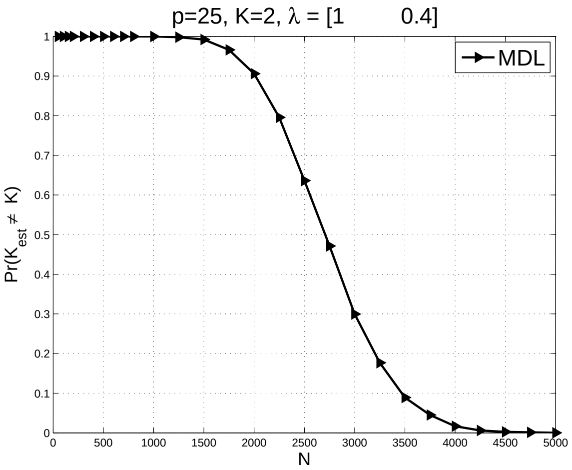
$$\hat{k}_{MDL} = \arg\min \quad -(p-k)n\left[\frac{1}{p-k}\log\left(\prod_{k+1}^{p}\ell_j\right) - \log\frac{\sum_{k+1}^{p}\ell_j}{p-k}\right] + \frac{1}{2}k(2p-k)\log n$$

$$\hat{k}_{AIC} = \arg\min \quad -(p-k)n\left[\frac{1}{p-k}\log\left(\prod_{k+1}^{p}\ell_j\right) - \log\frac{\sum_{k+1}^{p}\ell_j}{p-k}\right] + k(2p-k)$$

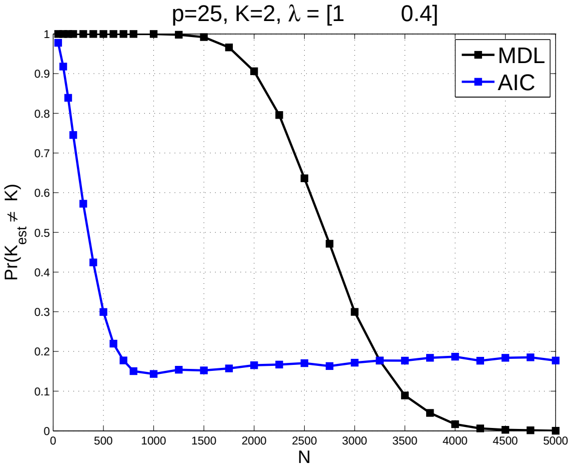
MDL estimator became the standard tool in signal processing.

WAX & KAILATH 1985.

Example: MDL consistent



Example: AIC not consistent



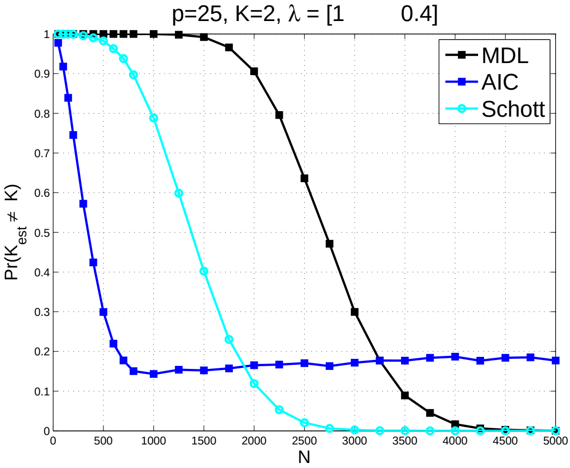
Some Modern Approaches:

Rao & Edelman (2007) and Schott (2006)

$$T = \frac{\sum \ell_j^2}{(\sum \ell_j)^2}$$

This statistic can be used for sphericity test for all values of p/n (Ledoit & Wolf).

Example:


p=25, K=2, λ = [1 0.4]
MDL
AIC
Schott
Pr(K_est ≠ K)
N
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
500
1000
1500
2000
2500
3000
3500
4000
4500
5000

Detection of Signals - Theoretical Questions

- Which Test Statistic to use ?
- Why does AIC overestimate number of signals ?
- Detection Performance ?
- Known vs. unknown noise level
- Non-parametric vs fully parametric setting.

KEY TOOLS

- Behavior of noise eigenvalues.

KEY TOOLS

- Behavior of noise eigenvalues.
- Behavior of signal eigenvalues in presence of noise.

KEY TOOLS

- Behavior of noise eigenvalues.
- Behavior of signal eigenvalues in presence of noise.
- Noise Estimate.

KEY TOOLS

- Behavior of noise eigenvalues.
- Behavior of signal eigenvalues in presence of noise.
- Noise Estimate.

TOOLS:

- RMT

- Matrix Perturbation Theory

Which Test Statistic ?

Consider the case of two (nearly) *simple* hypothesis

$$\mathcal{H}_0 : \mathbf{\Sigma} = \mathbf{I} \quad vs. \quad \mathcal{H}_1 : \mathbf{W}'\mathbf{\Sigma}\mathbf{W} = \mathbf{I} + diag(\lambda, 0, \ldots, 0)$$

with λ - known. What is unknown is the basis which makes $\mathbf{\Sigma}$ diagonal in $\mathcal{H}_1$.
Neyman-Pearson: optimal method is likelihood ratio test

$$\frac{p(\ell_1, \ldots, \ell_p | \mathcal{H}_1)}{p(\ell_1, \ldots, \ell_p | \mathcal{H}_0)} \gtrless C(\alpha)$$

Which Test Statistic ?

From multivariate analysis (Muirhead 78')

$$p(\ell_1, \ldots, \ell_p | \Sigma) = C_{n,p} \prod \ell_i^{(n-p-1)/2} \prod_{i<j} (\ell_i - \ell_j) \; {}_0F_0(-\frac{1}{2} nL, \Sigma^{-1})$$

${}_0F_0$ - hypergeometric function with matrix argument.
Key point: asymptotically in n, for p fixed,

$$\log \left(\frac{p(\ell_1, \ldots, \ell_p | \mathcal{H}_1)}{p(\ell_1, \ldots, \ell_p | \mathcal{H}_0)} \right) \approx n(\ell_1 - h(\lambda)) + O(\sum c_{1j}/(\ell_1 - \ell_j))$$

Asymptotically, should only look at *largest* eigenvalue !

Roy's Largest Eigenvalue Test

If $\ell_1 > \sigma^2 \cdot th(\alpha)$ accept $\mathcal{H}_1$ – signal is present.

$th(\alpha)$ - found by distribution of ℓ_1 under the null of no signals. Then, $S_n = 1/n \sum_i \mathbf{x}_i \mathbf{x}_i'$ is a scaled Wishart matrix.

Theorem: [Johansson 00', Johnstone 01',El-Karoui 07'] As $p, n \to \infty$

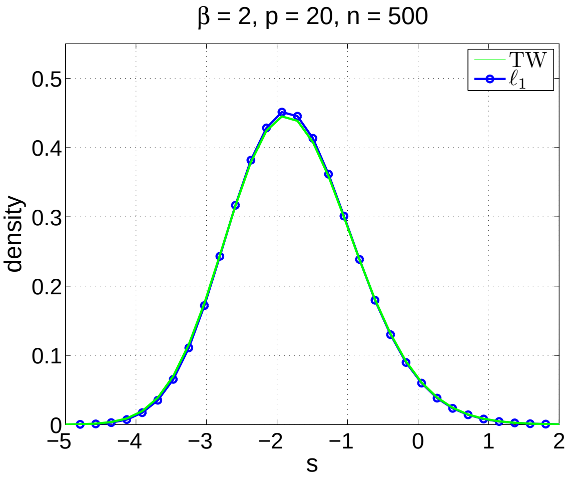
$$\boxed{\Pr\{\ell_1 < \sigma^2 (\mu_{n,p} + s\sigma_{n,p})\} \to TW_\beta(s)}$$

where TW_β - Tracy-Widom distribution of order β

$\beta = 1$ - real valued noise, $\beta = 2$ - complex valued noise.

For any confidence level α can invert TW distribution to obtain threshold $s(\alpha)$.

Largest Eigenvalue Distribution



Unknown Noise Level

Standard approach: replace σ^2 by its ML estimate

$$\hat{\sigma}^2 = \frac{1}{p}\sum_{j=1}^{p} \ell_j$$

GLRT:

$$U = \frac{\ell_1}{\frac{1}{p}\sum_j \ell_j} > th(\alpha)$$

This test statistic plays a role in:

- Signal Detection [Besson & Scharf 06', Kritchman & N. 08, Bianchi et al. 09']
- Two-way models of interaction [Johnson & Graybill, 72']
- Models for Quantum Information Channels.

Unknown Noise Level

Standard approach: replace σ^2 by its ML estimate

$$\hat{\sigma}^2 = \frac{1}{p}\sum_{j=1}^{p}\ell_j$$

GLRT:

$$U = \frac{\ell_1}{\frac{1}{p}\sum_j \ell_j} > th(\alpha)$$

This test statistic plays a role in:

- Signal Detection [Besson & Scharf 06', Kritchman & N. 08, Bianchi et al. 09']
- Two-way models of interaction [Johnson & Graybill, 72']
- Models for Quantum Information Channels.

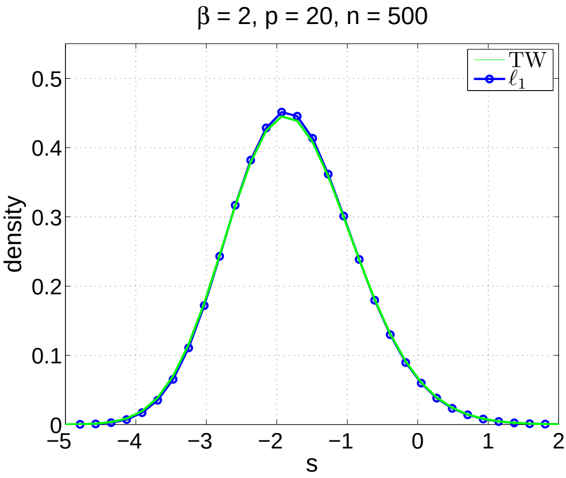
Unknown Noise Level

In principle, can use same threshold $th(\alpha)$, since:

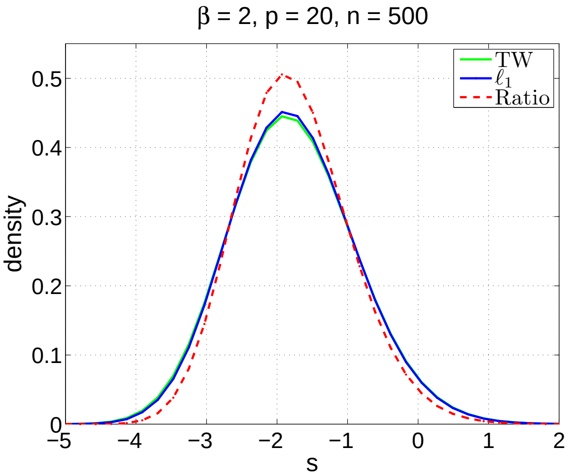
Theorem:[Bianchi et. al.] As $p, n \to \infty$

$$\Pr[U < \mu_{n,p} + s\sigma_{n,p}] \to TW_\beta(s)$$

Ratio Distribution



Ratio Distribution



Ratio Distribution

Definition:

$$F_{np}(s) = \Pr\left[\frac{\ell_1 - \mu_{np}}{\sigma_{np}} < s\right], \quad H_{np}(s) = \Pr\left[\frac{U - \mu_{np}}{\sigma_{np}} < s\right].$$

Ratio Distribution

Definition:

$$F_{np}(s) = \Pr\left[\frac{\ell_1 - \mu_{np}}{\sigma_{np}} < s\right], \quad H_{np}(s) = \Pr\left[\frac{U - \mu_{np}}{\sigma_{np}} < s\right].$$

Question: What is relation between F_{np} and H_{np} ?

Ratio Distribution

Definition:

$$F_{np}(s) = \Pr\left[\frac{\ell_1 - \mu_{np}}{\sigma_{np}} < s\right], \quad H_{np}(s) = \Pr\left[\frac{U - \mu_{np}}{\sigma_{np}} < s\right].$$

Question: What is relation between F_{np} and H_{np} ?

Key property: U and $T = \frac{1}{p}\sum_j \ell_j$ are independent.

Ratio Distribution

Since $\ell_1 = U \cdot T$,

$$\Pr[\ell_1 < x] = \int_0^{px} \Pr\left[U < \frac{x}{t}\right] p_T(t)dt$$

Ratio Distribution

Since $\ell_1 = U \cdot T$,

$$\Pr[\ell_1 < x] = \int_0^{px} \Pr\left[U < \frac{x}{t}\right] p_T(t)dt$$

L.H.S. $=$ approximately TW
R.H.S. $=$ convolution of required function with χ^2 density.

Ratio Distribution

Assumptions: In the joint limit $p, n \to \infty$ with $p/n \to c$, the following two conditions hold:

Ratio Distribution

Assumptions: In the joint limit $p, n \to \infty$ with $p/n \to c$, the following two conditions hold:

(i) uniformly in p and s, $H_{np}(s)$ is a smooth function with bounded third derivative, $|H_{np}'''(s)| < C$.

Ratio Distribution

Assumptions: In the joint limit $p, n \to \infty$ with $p/n \to c$, the following two conditions hold:

(i) uniformly in p and s, $H_{np}(s)$ is a smooth function with bounded third derivative, $|H'''_{np}(s)| < C$.

(ii)

$$|F'_{np}(s) - TW'_{\beta}(s)| \to 0 \qquad \text{and} \qquad |F''_{np}(s) - TW''_{\beta}(s)| \to 0 \quad (2)$$

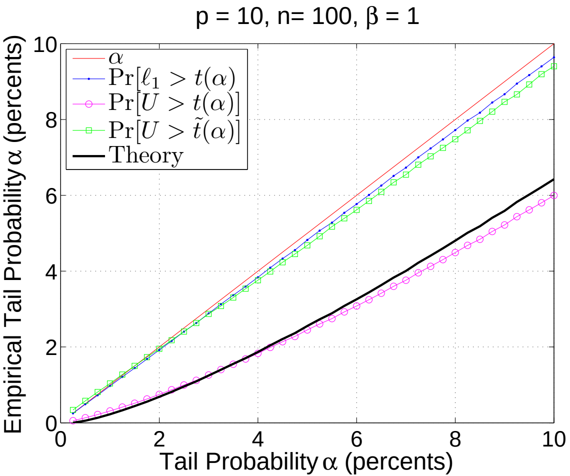
Ratio Distribution

Theorem: In the joint limit $p, n \to \infty$,

$$
\begin{aligned}
H_{np}(s) - TW_\beta(s) &= [F_{np}(s) - TW_\beta(s)] - \\
&\quad \frac{1}{2}\left(\frac{2}{\beta np}\right)\left(\frac{\mu_{np}}{\sigma_{np}}\right)^2 TW''_\beta(s) + o(p^{-2/3}).
\end{aligned}
$$

[N., to appear in JMVA]

Ratio Distribution Tail Probabilities



Ratio Distribution Tail Probabilities

When is correction term *small* w.r.t. $1 - TW_\beta(s)$?

Ratio Distribution Tail Probabilities

When is correction term *small* w.r.t. $1 - TW_\beta(s)$?

Example: $n \gg p$, $s = -0.2325$, where

$$1 - TW_2(s) \approx 5\%$$

Then

$$\frac{|TW_2''(s)|}{1 - TW_2(s)} \approx 7 \quad \text{and} \quad \frac{1}{np}\left(\frac{\mu_{np}}{\sigma_{np}}\right)^2 \approx 1/p^{2/3}.$$

Ratio Distribution Tail Probabilities

When is correction term *small* w.r.t. $1 - TW_\beta(s)$?

Example: $n \gg p$, $s = -0.2325$, where

$$1 - TW_2(s) \approx 5\%$$

Then

$$\frac{|TW_2''(s)|}{1 - TW_2(s)} \approx 7 \quad \text{and} \quad \frac{1}{np}\left(\frac{\mu_{np}}{\sigma_{np}}\right)^2 \approx 1/p^{2/3}.$$

Hence, for a 10% relative error,

$$\frac{1}{2}\frac{2}{\beta np}\left(\frac{\mu_{np}}{\sigma_{np}}\right)^2 \frac{|TW_2''(s)|}{1 - TW_2(s)} \leq 0.1$$

we need $p \gtrsim (35)^{3/2} \approx 200$.

Detection of Signals - Theoretical Questions

- Which Test Statistic to use ?
- **Why does AIC overestimate number of signals**
- Detection Performance ?
- Known vs. unknown noise level
- Non-parametric vs fully parametric setting.

Overestimation Probability of AIC (no signals)

AIC overestimates number of signals when

$$-\ln \mathcal{L}_0 > -\ln \mathcal{L}_1 + \frac{2p-1}{n}$$

where

$$\mathcal{L}_k = \frac{\prod_{i>k} \ell_i}{\left(\frac{\sum_{i>k} \ell_i}{p-k}\right)}$$

Overestimation Probability of AIC

Let ξ be solution of

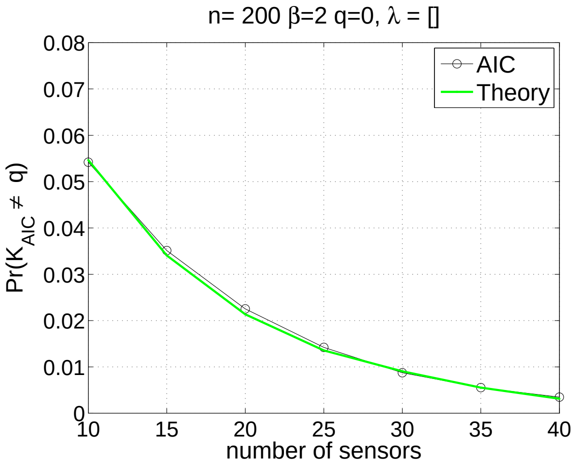
$$-\ln(1+\xi) - (p-1)\ln(1 - \frac{\xi}{p-1}) = \frac{2p-1}{n}$$

Lemma:

$$\Pr[k_{\text{AIC}} > 0] = \Pr[U > 1 + \xi] + O\left(\frac{1}{n}\right)$$

[N., IEEE-TSP 10']

AIC Overestimation Probability



AIC Overestimation Probability

Since

$$\xi = 2\sqrt{\frac{p-1/2}{n}}\sqrt{1-\frac{1}{p}}\left(1+O\left(\frac{1}{n^{1/2}}\right)\right)$$

AIC penalty is not sufficiently strong.

AIC Overestimation Probability

Since

$$\xi = 2\sqrt{\frac{p-1/2}{n}}\sqrt{1-\frac{1}{p}}\left(1+O\left(\frac{1}{n^{1/2}}\right)\right)$$

AIC penalty is not sufficiently strong.

MDL penalty too large since contains a $\ln n$ factor.

AIC Overestimation Probability

Since

$$\xi = 2\sqrt{\frac{p-1/2}{n}}\sqrt{1-\frac{1}{p}}\left(1+O\left(\frac{1}{n^{1/2}}\right)\right)$$

AIC penalty is not sufficiently strong.

MDL penalty too large since contains a $\ln n$ factor.

Legacy System:

$$\hat{k}_{AIC} = \arg\min \qquad -\ln \mathcal{L}_k + \frac{2Ck}{n}\left(p+1-\frac{k+1}{2}\right)$$

If $C = 1$ penalty comparable to original AIC.

Modified AIC

Theorem: Modified AIC estimator with $C = 2$ has negligible overestimation probability, which for large n is exponential small in p.

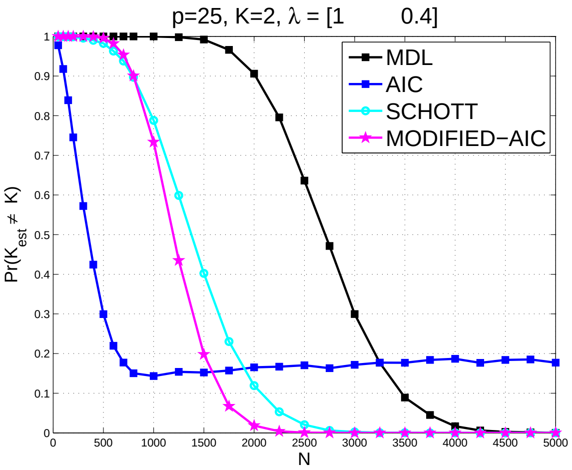
Modified AIC

Theorem: Modified AIC estimator with $C = 2$ has negligible overestimation probability, which for large n is exponential small in p.

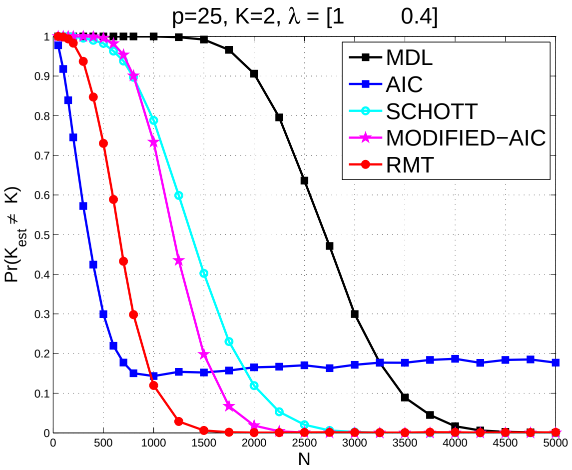
Basic idea: use non-asymptotic bound of Ledoux, valid for all p, n (complex)

$$\Pr[\ell_1 > (1 + \sqrt{p/n})^2 + \epsilon] \leq e^{-nJ_{LAG}(\epsilon)}$$

Modified AIC



Example:


p=25, K=2, λ = [1 0.4]
MDL
AIC
SCHOTT
MODIFIED−AIC
RMT
$Pr(K_{est} \neq K)$
N
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
500
1000
1500
2000
2500
3000
3500
4000
4500
5000

Detection of Signals - Theoretical Questions

- Which Test Statistic to use ?
- Why does AIC overestimate number of signals ?
- **Detection Performance ?**
- Known vs. unknown noise level
- Non-parametric vs fully parametric setting.

Detection Performance

Depends on behavior of largest signal eigenvalue in presence of noise. With p fixed, $n \to \infty$,

$$P_d^{Roy} \approx Q\left[\sqrt{n}\left(\frac{th(\alpha)}{1+\lambda/\sigma^2} - \frac{p-1}{N\lambda/\sigma^2} - 1\right)\right]$$

Detection Performance

Depends on behavior of largest signal eigenvalue in presence of noise. With p fixed, $n \to \infty$,

$$P_d^{Roy} \approx Q\left[\sqrt{n}\left(\frac{th(\alpha)}{1+\lambda/\sigma^2} - \frac{p-1}{N\lambda/\sigma^2} - 1\right)\right]$$

$$P_d^{GLRT} \approx Q\left[\sqrt{n}\left(\tilde{th}(\alpha)(\frac{1}{1+\lambda/\sigma^2} - \frac{1}{n\lambda/\sigma^2}) - \frac{p-1}{N\lambda/\sigma^2} - 1\right)\right]$$

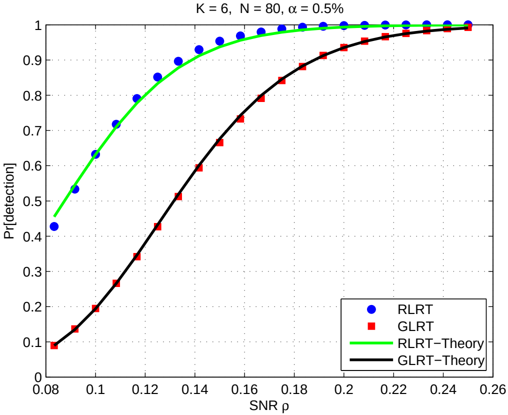
where $Q(z) = \frac{1}{\sqrt{2\pi}} \int_z^\infty e^{-x^2/2} dx$,

$$\tilde{th}(\alpha) = \frac{p-1}{p - th_U(\alpha)} th_U(\alpha) > th(\alpha)$$

[Kritchman & N. 09']
[N, Penna, Garello, submitted]

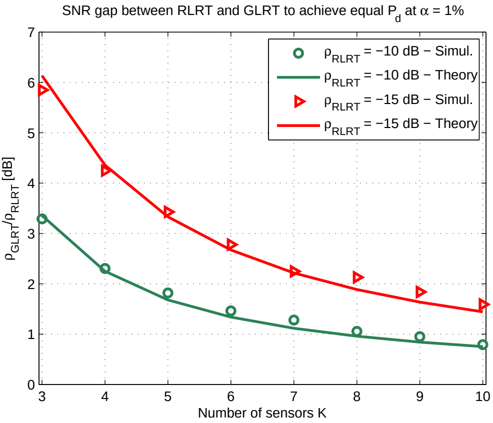
Known vs. Unknown Noise Level

Difference can be large:



Known vs. Unknown Noise Level

Difference can be large (several dB)



Detection of Signals - Theoretical Questions

- Which Test Statistic to use ?
- Why does AIC overestimate number of signals ?
- Detection Performance ?
- Known vs. unknown noise level
- Non-parametric vs fully parametric setting.

Non-Parametric vs. Parametric Detection

In signal array processing, the array manifold is typically known.
For Uniform Linear Array

$$\mathbf{a}(\theta) = \left[1 e^{i\pi \sin\theta} e^{2i\pi \sin(\theta)} \ldots e^{(p-1)i\pi \sin\theta} \right]$$

Non-Parametric vs. Parametric Detection

In signal array processing, the array manifold is typically known.
For Uniform Linear Array

$$\mathbf{a}(\theta) = \left[1 e^{i\pi \sin\theta} e^{2i\pi \sin(\theta)} \ldots e^{(p-1)i\pi \sin\theta}\right]$$

Instead of largest eigenvalue of sample covariance matrix, maximal correlation of $\mathbf{x}$ with $\mathbf{a}(\theta)$

Non-Parametric vs. Parametric Detection

In signal array processing, the array manifold is typically known.
For Uniform Linear Array

$$\mathbf{a}(\theta) = \left[1 e^{i\pi \sin\theta} e^{2i\pi \sin(\theta)} \ldots e^{(p-1)i\pi \sin\theta} \right]$$

Instead of largest eigenvalue of sample covariance matrix, maximal correlation of $\mathbf{x}$ with $\mathbf{a}(\theta)$

Theorem: Parametric Signal Detection requires

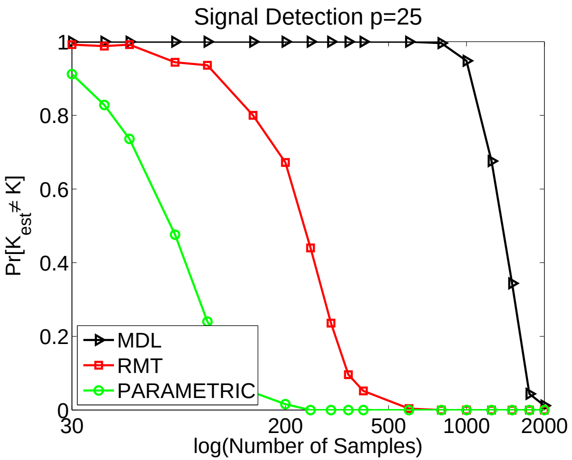
$$SNR \gg \sqrt{\frac{2 \ln p}{n}}$$

instead of

$$SNR \gg \sqrt{\frac{p}{n}}$$

[Arkind & N., 10']

Example:


Signal Detection p=25
Pr[K_est ≠ K]
log(Number of Samples)
0
0.2
0.4
0.6
0.8
1
30
200
500
1000
2000
MDL
RMT
PARAMETRIC

Summary and Open Questions

Open Questions:

- Detection in unknown noise environments
- Non-Gaussian signals
- Sparse covariance matrices

http://www.wisdom.weizmann.ac.il/∼nadler/

Summary and Open Questions

Open Questions:

- Detection in unknown noise environments
- Non-Gaussian signals
- Sparse covariance matrices

http://www.wisdom.weizmann.ac.il/∼nadler/

C'est Tou / Merci Beaucoup