Aalto University
School of Science
and Technology

On Demmel Condition Number Distributions with Applications in Telecommunications

Lu Wei and Olav Tirkkonen

Aalto University, Finland
Joint work with Matthew R. McKay, HKUST, Hong Kong

12.Oct.2010

Outline

Demmel Condition Number

- Definition
- Existing results

Derivations for DCN Distributions

- General framework
- Exact distribution
- Asymptotic distribution

Applications in Wireless Communications

- Adaptive transmission
- Adaptive detection

Definition

- Define a $K \times N$ dimension matrix $\mathbf{X}$ with independent and identically distributed (i.i.d) complex Gaussian entries, each with zero mean and unit variance.
- The $K \times K$ Hermitian matrix $\mathbf{R} = \mathbf{X}\mathbf{X}^\dagger$ follows a complex Wishart distribution with N degree of freedom (d.o.f).
- We denote the ordered eigenvalues of $\mathbf{R}$ as $\lambda_1 > \lambda_2 > ... > \lambda_K > 0$, and the trace of $\mathbf{R}$ as $T = \text{tr}\{\mathbf{R}\} = ||\mathbf{X}||_F^2 = \sum_{i=1}^{K} \lambda_i$, where $|| \cdot ||_F$ is the Frobenius norm.

Definition

- Define a $K \times N$ dimension matrix $\mathbf{X}$ with independent and identically distributed (i.i.d) complex Gaussian entries, each with zero mean and unit variance.
- The $K \times K$ Hermitian matrix $\mathbf{R} = \mathbf{X}\mathbf{X}^\dagger$ follows a complex Wishart distribution with N degree of freedom (d.o.f).
- We denote the ordered eigenvalues of $\mathbf{R}$ as $\lambda_1 > \lambda_2 > ... > \lambda_K > 0$, and the trace of $\mathbf{R}$ as $T = \text{tr}\{\mathbf{R}\} = ||\mathbf{X}||_F^2 = \sum_{i=1}^{K} \lambda_i$, where $|| \cdot ||_F$ is the Frobenius norm.
- The Demmel Condition Number (DCN) of $\mathbf{R}$ is defined as the ratio of its trace to its smallest eigenvalue λ_K,

$$X := \frac{\sum_{i=1}^{K} \lambda_i}{\lambda_K} = \frac{T}{\lambda_K}, \tag{1}$$

where $x \in [K, \infty]$.

Existing results

- Limited results on the DCN distribution exist in the literature.
- A. Edelman, "On the distribution of a scaled condition number," *Math. Comp., vol. 58, pp. 185-190, 1992.*
  - *Exact DCN distributions for the special case $K = N$ (both real and complex cases).*
  - *Mainly based on the fact that λ_K has tractable expressions when $K = N$ (e.g. exponentially distributed in complex case).*
  - *Using an equality (A. W. Davis,* 1972*) between Laplace transforms of PDFs of X and λ_K.*

Existing results

- M. Matthaiou, M. R. McKay, P. J. Smith, and J. A. Nossek, "On the condition number distribution of complex Wishart matrices," *IEEE Tran. Commun., vol. 58, no. 6, pp. 1705-1711, Jun. 2010.*
  - *Exact DCN distributions for $K = 2$ with arbitrary N.*
  - *Established through standard condition number distribution ($\frac{\lambda_1+\lambda_2}{\lambda_2} = 1 + \frac{\lambda_1}{\lambda_2}$).*
- *Above two results are exact. No asymptotic results w.r.t. matrix dimension are available.*

Existing results

- M. Matthaiou, M. R. McKay, P. J. Smith, and J. A. Nossek, "On the condition number distribution of complex Wishart matrices," *IEEE Tran. Commun., vol. 58, no. 6, pp. 1705-1711, Jun. 2010.*
  - *Exact DCN distributions for* $K = 2$ *with arbitrary* N.
  - *Established through standard condition number distribution* $(\frac{\lambda_1+\lambda_2}{\lambda_2} = 1 + \frac{\lambda_1}{\lambda_2})$.
- *Above two results are exact. No asymptotic results w.r.t. matrix dimension are available.*
- *In this work, both exact and asymptotic DCN distributions for arbitrary* K *and* N *are derived.*

General framework

- Intractable correlation between T and λ_K exists.

General framework

- Intractable correlation between T and λ_K exists.
- But, it can be verified (O. Besson, 2006) that $Y := \lambda_K / T$ and T are independent.

General framework

- Intractable correlation between T and λ_K exists.
- But, it can be verified (O. Besson, 2006) that $Y := \lambda_K / T$ and T are independent.
- Thus, λ_K equals the product of the independent r.v Y and T. Define $f(x)$, $g(x)$ and $h(x)$ as the PDFs of λ_K, T and Y respectively.

General framework

- Intractable correlation between T and λ_K exists.
- But, it can be verified (O. Besson, 2006) that $Y := \lambda_K / T$ and T are independent.
- Thus, λ_K equals the product of the independent r.v Y and T. Define $f(x)$, $g(x)$ and $h(x)$ as the PDFs of λ_K, T and Y respectively.
- By this independence, it holds that

$$M_z[f(x)] = M_z[g(x)] M_z[h(x)], \tag{2}$$

where $M_z[\cdot]$ denotes Mellin transform.

General framework

- By Mellin inversion integral, the distribution of $h(x)$ can be uniquely determined by

$$h(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} x^{-z} \frac{M_z[f(x)]}{M_z[g(x)]} \mathrm{d}z. \tag{3}$$

- A transform from Y to $1/Y$ yields the desired DCN PDF.

General framework

- By Mellin inversion integral, the distribution of $h(x)$ can be uniquely determined by

$$h(x) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} x^{-z}\frac{M_z[f(x)]}{M_z[g(x)]}\mathrm{d}z. \tag{3}$$

- A transform from Y to $1/Y$ yields the desired DCN PDF.
- Merits of this framework:
  - Correlation between λ_K and T is implicitly taken into account by the product of Mellin transforms (2).
  - Mellin inversion integral (3) can be easily evaluated by the residue theorem.

General framework

- By Mellin inversion integral, the distribution of $h(x)$ can be uniquely determined by

$$h(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} x^{-z} \frac{M_z[f(x)]}{M_z[g(x)]} \mathrm{d}z. \tag{3}$$

- A transform from Y to $1/Y$ yields the desired DCN PDF.
- Merits of this framework:
  - Correlation between λ_K and T is implicitly taken into account by the product of Mellin transforms (2).
  - Mellin inversion integral (3) can be easily evaluated by the residue theorem.
- This framework provides the possibility to obtain both exact and asymptotic DCN distributions.

Exact distribution

- C. S. Park and K. B. Lee "Statistical multimode transmit antenna selection for limited feedback MIMO systems," *IEEE Tran. Wireless Commun., vol. 7, no. 11, pp. 4432-4438, Nov. 2008.*
  - *PDF of λ_K represented as a weighted sum of polynomials as*

$$f(x) = e^{-Kx} \sum_{n=N-K}^{(N-K)K} c_n^{(N,K)} x^n. \tag{4}$$

  - *Coefficients $c_n^{(N,K)}$ is determined by the symmetry of the integral representation of λ_K (A. Edelman,* 1989*).*

Determining the coefficients $c_n^{(N,K)}$

- Define

$$I_n(m) := \sum_{k=0}^{n} \binom{n}{k} (m+n-k)!x^k. \tag{5}$$

- $K = 2$, PDF of λ_K is

$$c_2 e^{-2x} x^{N-2} I_{N-2}(2). \tag{6}$$

Determining the coefficients $c_n^{(N,K)}$

- Define

$$I_n(m) := \sum_{k=0}^{n} \binom{n}{k} (m+n-k)! x^k. \tag{5}$$

- $K = 2$, PDF of λ_K is

$$c_2 e^{-2x} x^{N-2} I_{N-2}(2). \tag{6}$$

- $K = 3$, PDF of λ_K is

$$c_3 e^{-3x} x^{N-3} [I_{N-3}(4) I_{N-3}(2) - (I_{N-3}(3))^2]. \tag{7}$$

Determining the coefficients $c_n^{(N,K)}$

- Define

$$I_n(m) := \sum_{k=0}^{n} \binom{n}{k} (m+n-k)!x^k. \qquad (5)$$

- $K = 2$, PDF of λ_K is

$$c_2 e^{-2x} x^{N-2} I_{N-2}(2). \qquad (6)$$

- $K = 3$, PDF of λ_K is

$$c_3 e^{-3x} x^{N-3} [I_{N-3}(4) I_{N-3}(2) - (I_{N-3}(3))^2]. \qquad (7)$$

Determining the coefficients $c_n^{(N,K)}$

- $K = 4$, PDF of λ_K is

$$c_4 e^{-4x} x^{N-4}[I_{N-4}(6) I_{N-4}(4) I_{N-4}(2) - I_{N-4}(6)(I_{N-4}(3))^2 \quad (8)$$
$$+2 I_{N-4}(5) I_{N-4}(4) I_{N-4}(3) - (I_{N-4}(5))^2 I_{N-4}(2) - (I_{N-4}(4))^2].$$

Determining the coefficients $c_n^{(N,K)}$

- $K = 4$, PDF of λ_K is

$$c_4 e^{-4x} x^{N-4}[I_{N-4}(6)I_{N-4}(4)I_{N-4}(2) - I_{N-4}(6)(I_{N-4}(3))^2 \quad (8)$$
$$+2I_{N-4}(5)I_{N-4}(4)I_{N-4}(3) - (I_{N-4}(5))^2 I_{N-4}(2) - (I_{N-4}(4))^2].$$

- Note:
  - After some basic manipulations, the expressions for coefficients of x can be obtained.
  - Although tedious, coefficients for arbitrary K can be similarly calculated.

Exact distribution

- Using the closed-form expression for PDF of λ_K, the developed framework can be applied.

Exact distribution

- Using the closed-form expression for PDF of λ_K, the developed framework can be applied.
- We first calculate,

$$
\begin{aligned}
M_z[f(x)] &= \sum_{n=N-K}^{(N-K)K} c_n^{(N,K)} \frac{\Gamma(z+n)}{K^{z+n}}, \\
M_z[g(x)] &= \frac{1}{\Gamma(m/2)} \Gamma(z + \frac{m}{2} - 1), \quad (m = 2KN).
\end{aligned}
$$

Exact distribution

- Using the closed-form expression for PDF of λ_K, the developed framework can be applied.
- We first calculate,

$$\begin{aligned} M_z[f(x)] &= \sum_{n=N-K}^{(N-K)K} c_n^{(N,K)} \frac{\Gamma(z+n)}{K^{z+n}}, \\ M_z[g(x)] &= \frac{1}{\Gamma(m/2)} \Gamma(z + \frac{m}{2} - 1), \quad (m = 2KN). \end{aligned}$$

- Using the residue theorem, $h(x)$ is uniquely determined to be

$$h(x) = \frac{\Gamma(m/2)}{(1-Kx)^{2-m/2}} \sum_{n=N-K}^{(N-K)K} \frac{c_n^{(N,K)}}{\Gamma(m/2-n-1)} \left(\frac{x}{1-Kx}\right)^n. \tag{9}$$

Exact distribution

- By a simple transform, PDF of DCN is obtained as,

$$d(x) = \frac{\Gamma(m/2)x^{-m/2}}{(x-K)^{2-m/2}} \sum_{n=N-K}^{(N-K)K} \frac{c_n^{(N,K)}}{\Gamma(m/2-n-1)}(x-K)^{-n}. \tag{10}$$

Exact distribution

- By a simple transform, PDF of DCN is obtained as,

$$d(x) = \frac{\Gamma(m/2)x^{-m/2}}{(x-K)^{2-m/2}} \sum_{n=N-K}^{(N-K)K} \frac{c_n^{(N,K)}}{\Gamma(m/2-n-1)}(x-K)^{-n}. \tag{10}$$

- Then CDF of DCN is calculated to be,

$$D(y) = \Gamma(\frac{m}{2}) \sum_{n=N-K}^{(N-K)K} \frac{K^{-n-1}c_n^{(N,K)}}{\Gamma(m/2-n-1)} \left(B(a,b) - B_{\frac{K}{y}}(a,b)\right) \tag{11}$$

$B_x(a,b)$ and $B(a,b)$ are incomplete and complete Beta function respectively and $a = n+1$, $b = \frac{m}{2} - n - 1$.

Special cases

- Here we check the derived result with some known special cases.

Special cases

- Here we check the derived result with some known special cases.
- $K = N$ (A. Edelman, 1992)
  - The only coefficient left is $c_0^{(K,K)} = K$.
  - Inserting this coefficient into the derived PDF, $d(x)$ simplifies to

$$d(x) = K(K^2 - 1)x^{-K^2}(x - K)^{K^2-2}. \tag{12}$$

  - Agrees with the known result.

Special cases

- $K = 2$, with arbitrary N (M. Matthaiou, 2010)
  - The coefficient in this case is

$$c_n^{(N,2)} = \frac{\Gamma(2N - n - 1)}{\Gamma(N)\Gamma(n - N + 3)\Gamma(2N - n - 3)}. \tag{13}$$

  - Inserting $c_n^{(N,2)}$ into the derived PDF, $d(x)$ simplifies to

$$d(x) = \frac{\Gamma(2N)}{\Gamma(N)\Gamma(N - 1)}(x - 2)^2 x^{-2N}(x - 1)^{N-2}. \tag{14}$$

  - In agreement with the known result.

One numerical example

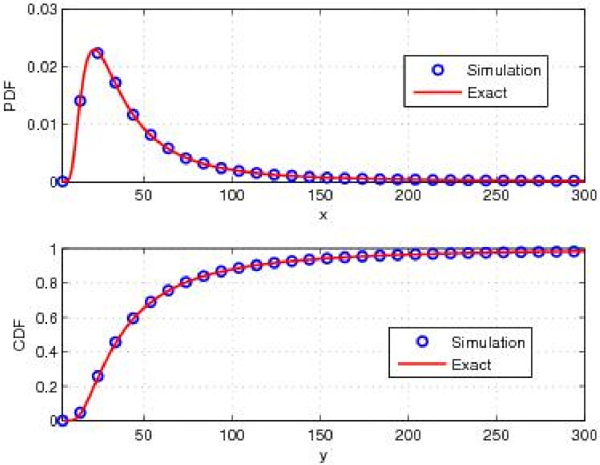
- $K = 4$, $N = 5$.
- $d(x)$ is calculated to be

$$3420(x-4)^{14}(x^3+5x^2-20x+4)x^{-20}. \tag{15}$$

- $D(y)$ is calculated to be

$$\begin{aligned} &1-213.75B_{\frac{4}{y}}(2,18)-908.438B_{\frac{4}{y}}(3,17) \\ &-908.438B_{\frac{4}{y}}(4,16)-227.109B_{\frac{4}{y}}(5,15). \end{aligned} \tag{16}$$

One numerical example: $K = 4$, $N = 5$



Asymptotic distribution

- Motivation:
  - Determining the coefficients may appear a problem for large dimensional matrices.
  - We would like to gain insight into the behavior of DCN distribution when the dimension K, N are large.

Asymptotic distribution

- Motivation:
  - Determining the coefficients may appear a problem for large dimensional matrices.
  - We would like to gain insight into the behavior of DCN distribution when the dimension K, N are large.
- We derive a closed-form asymptotic DCN distribution, which circumvents the need to calculate the coefficients.
- The asymptotic result falls in the developed Mellin transform framework as well.

An asymptotic result on λ_k distribution

- For λ_k, there exists sequences $a(K, N)$ and $b(K, N)$ such that the distribution of the random variable

$$\Lambda_K = \frac{\lambda_K - a(K, N)}{b(K, N)} \tag{17}$$

  converges to the Tracy-Widom distribution of order two (O. N. Feldheim, 2010), denoted as F_{TW2}.
- This result provides an approximation to λ_k for large K and N,

$$F(x) \approx F_{\mathrm{TW2}}\left(\frac{x - a(K, N)}{b(K, N)}\right). \tag{18}$$

An asymptotic result on λ_k distribution

- For λ_k, there exists sequences $a(K, N)$ and $b(K, N)$ such that the distribution of the random variable

$$\Lambda_K = \frac{\lambda_K - a(K, N)}{b(K, N)} \tag{17}$$

  converges to the Tracy-Widom distribution of order two (O. N. Feldheim, 2010), denoted as F_{TW2}.
- This result provides an approximation to λ_k for large K and N,

$$F(x) \approx F_{\text{TW2}} \left(\frac{x - a(K, N)}{b(K, N)} \right). \tag{18}$$

- Numerical burden to calculate $F_{\text{TW2}}(\cdot)$, simpler closed-form approximation is more desirable.

Gamma approximation to λ_k distribution

- It was stated in (A. Edelman, 2005) that λ_k can be well approximated by a Gamma distribution.

Gamma approximation to λ_k distribution

- It was stated in (A. Edelman, 2005) that λ_k can be well approximated by a Gamma distribution.
- Motivated by this, we propose a Gamma approximation by calculating the first two asymptotic moments via Tracy-Widom distribution.
  - $E[\lambda_K] = a(K,N) + b(K,N)E[\Lambda_K]$.
  - $V[\lambda_K] = \big(b(K,N)\big)^2 V[\Lambda_K]$.
- Convergence in distribution implies

$$E[\Lambda_K] \to E[x_{\text{TW2}}] = -1.7711, \tag{19}$$

$$V[\Lambda_K] \to V[x_{\text{TW2}}] = 0.8132. \tag{20}$$

Gamma approximation to λ_k distribution

- For a Gamma distribution with parameters θ and k, by matching the two moments of λ_k, θ and k is obtained as

$$k = \frac{\big(a(K,N) + b(K,N)E[x_{TW2}]\big)^2}{\big(b(K,N)\big)^2 V[x_{TW2}]}, \quad (21)$$

$$\theta = \frac{\big(b(K,N)\big)^2 V[x_{TW2}]}{a(K,N) + b(K,N)E[x_{TW2}]}. \quad (22)$$

Asymptotic distribution

- Using the closed-form asymptotic λ_K distribution, the developed framework can be applied.

Asymptotic distribution

- Using the closed-form asymptotic λ_K distribution, the developed framework can be applied.
- We first calculate

$$M_z[f(x)] = \frac{\theta^{z-1}}{\Gamma(k)}\Gamma(z+k-1). \qquad (23)$$

Asymptotic distribution

- Using the closed-form asymptotic λ_K distribution, the developed framework can be applied.
- We first calculate

$$M_z[f(x)] = \frac{\theta^{z-1}}{\Gamma(k)}\Gamma(z+k-1). \tag{23}$$

- By the residue theorem and a variable transform, the PDF of asymptotic DCN is calculated as

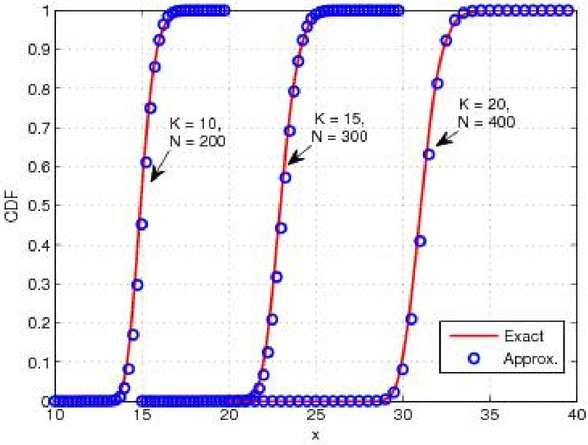
$$d(x) = c_1 x^{-m/2}(\theta x - 1)^{m/2-k-1}. \tag{24}$$

- Then, CDF of asymptotic DCN is calculated as

$$D(y) = c_2\big(C(K) - C(y)\big), \tag{25}$$

$C(x) = {}_2F_1(k, 1+k-\frac{m}{2}; k+1; \frac{1}{\theta x})x^{-k}$.

Numerical results



Application

- The $K \times N$ dimension matrix $\mathbf{X}$ models the MIMO communication channels.

Application

- The $K \times N$ dimension matrix $\mathbf{X}$ models the MIMO communication channels.
- Performance analysis and design of MIMO techniques relies on the statistical properties of the random MIMO channels.

Application

- The $K \times N$ dimension matrix $\mathbf{X}$ models the MIMO communication channels.
- Performance analysis and design of MIMO techniques relies on the statistical properties of the random MIMO channels.
- DCN reflects the eigenvalue spread of the random MIMO channel – indicates multipath richness for a given channel realization.

Application

- The $K \times N$ dimension matrix $\mathbf{X}$ models the MIMO communication channels.
- Performance analysis and design of MIMO techniques relies on the statistical properties of the random MIMO channels.
- DCN reflects the eigenvalue spread of the random MIMO channel – indicates multipath richness for a given channel realization.
- Using this fact, several MIMO transmit and receive schemes can be proposed.

Application: adaptive transmission

- Adaptive transmission can be achieved based on the Demmel condition number.

Application: adaptive transmission

- Adaptive transmission can be achieved based on the Demmel condition number.
- Transmission rate and reliability trade-off:
  - Spatial multiplexing: high data rate, no diversity.
  - Transmit diversity: lower data rate, possibility to achieve full diversity.

Application: adaptive transmission

- Adaptive transmission can be achieved based on the Demmel condition number.
- Transmission rate and reliability trade-off:
  - Spatial multiplexing: high data rate, no diversity.
  - Transmit diversity: lower data rate, possibility to achieve full diversity.
- Transmitter needs feedback information from receiver.

Application: adaptive transmission

- Adaptive transmission can be achieved based on the Demmel condition number.
- Transmission rate and reliability trade-off:
  - Spatial multiplexing: high data rate, no diversity.
  - Transmit diversity: lower data rate, possibility to achieve full diversity.
- Transmitter needs feedback information from receiver.
- Adaptive transmission switches between the two schemes depending on the instantaneous DCN.

Application: adaptive transmission

- Adaptive transmission can be achieved based on the Demmel condition number.
- Transmission rate and reliability trade-off:
  - Spatial multiplexing: high data rate, no diversity.
  - Transmit diversity: lower data rate, possibility to achieve full diversity.
- Transmitter needs feedback information from receiver.
- Adaptive transmission switches between the two schemes depending on the instantaneous DCN.
- Combining the benefits of the two transmission methods, switching is based hard decision.

Application: adaptive detection

- Adaptive detection is possible using the Demmel condition number.

Application: adaptive detection

- Adaptive detection is possible using the Demmel condition number.
- Detection performance and complexity trade-off:
  - Maximum likelihood detection: optimum with high complexity.
  - Zero forcing detection: sub-optimum with low complexity.

Application: adaptive detection

- Adaptive detection is possible using the Demmel condition number.
- Detection performance and complexity trade-off:
  - Maximum likelihood detection: optimum with high complexity.
  - Zero forcing detection: sub-optimum with low complexity.
- Adaptive detector switches between these two detection algorithms depending on the instantaneous DCN.

Application: adaptive detection

- Adaptive detection is possible using the Demmel condition number.
- Detection performance and complexity trade-off:
  - Maximum likelihood detection: optimum with high complexity.
  - Zero forcing detection: sub-optimum with low complexity.
- Adaptive detector switches between these two detection algorithms depending on the instantaneous DCN.
- Combining the benefits of both detectors, switching is based hard decision.

Thank you!