On the Role of the Hypergeometric Functions of Matrix Argument in the Age of Electronic Communications

Kenneth Gross (University of Vermont)

and

Donald Richards (Penn State University)

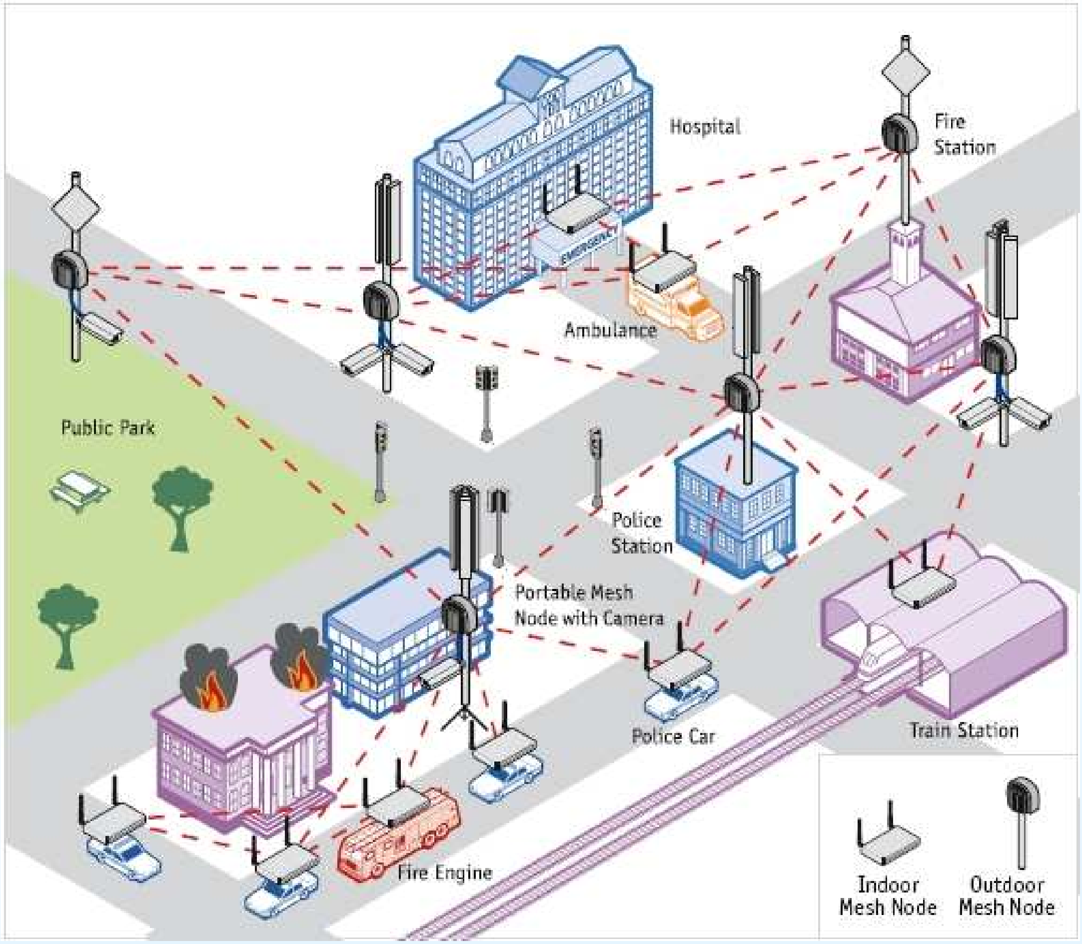
A 1989 paper

Gross and Richards, “Total positivity, spherical series, and hypergeometric functions of matrix argument,” Journal of Approximation Theory, 59 (1989), 224–246

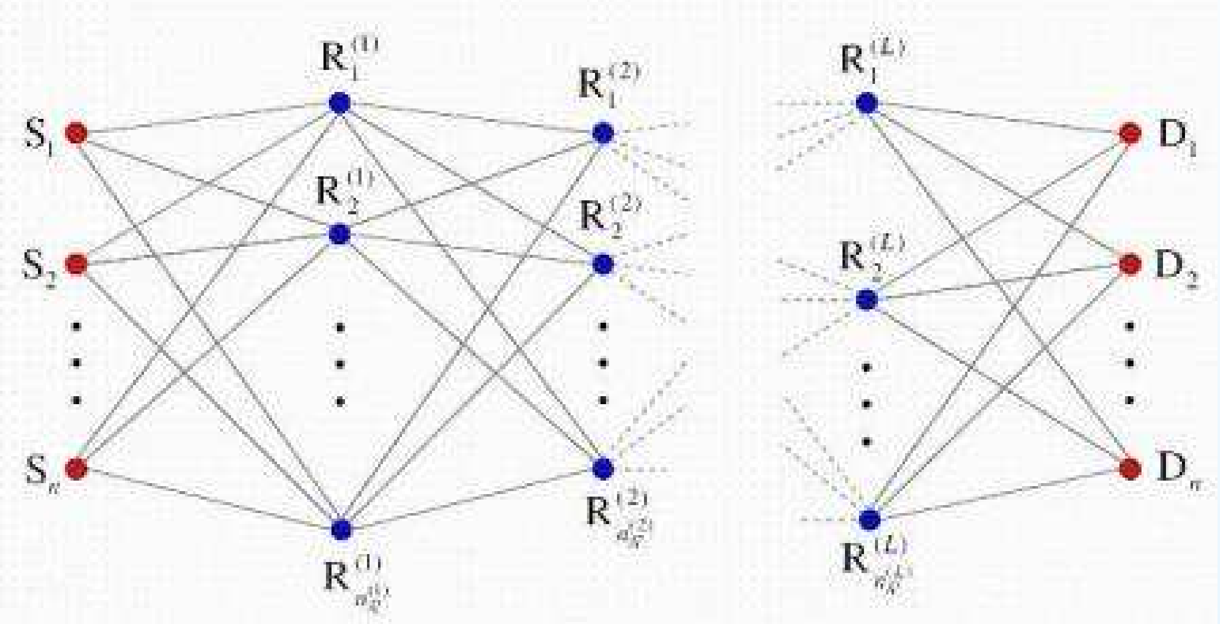
Google Scholar: 73 citations

Large numbers of citations by papers on wireless communications

Multiple-Input Multiple-Output (MIMO) communications system



Source: http://blog.wlanmall.com/wp-content/uploads/2010/07/publicsafety_diag.jpg



Source: http://www.nari.ee.ethz.ch/wireless/research/projects/figures/MimoMultiHop.jpg

Early difficulties in wireless transmission: Signal degradation due to buildings, terrain, etc.

Naturally occurring “noise” also caused signals to fade in transmission

The use of multiple antennas led to substantial increases in information transmission

MIMO system: The use of multiple antennas for transmitting and receiving wireless communications signals

The advantages:

Increased data throughput
Longer link range
No need for additional bandwith or transmit power
Higher spectral efficiency (more bits/second/hertz of bandwidth)
Reduced fading

A major development in wireless communications systems

http://en.wikipedia.org/wiki/MIMO

Signal transmission

A signal consists of a phase, θ, and an amplitude, r

Each signal is represented by a complex number, $re^{\mathrm{i}\theta}$

n_t: The number of transmitting antennas

n_r: The number of receiving antennas

$\boldsymbol{x}$: The $n_t \times 1$ vector of signals sent by the transmitting antennas

$\boldsymbol{y}$: The $n_r \times 1$ vector of signals arriving at the receiving antennas

Both $\boldsymbol{x}$ and $\boldsymbol{y}$ are vectors of complex numbers

We assume, w.l.o.g., that $n_r \geq n_t$

Channel matrix: A $n_r \times n_t$ complex random matrix, $\boldsymbol{G}$, relating $\boldsymbol{x}$ and $\boldsymbol{y}$

$\boldsymbol{\varepsilon}$: System noise, a $n_r \times 1$ complex random vector

A simple model for a MIMO system:

$$\boldsymbol{y} = \boldsymbol{G}\boldsymbol{x} + \boldsymbol{\varepsilon}$$

The entries of $\boldsymbol{G}$, $\boldsymbol{x}$, and $\boldsymbol{\varepsilon}$ are mutually independent, identically distributed random variables, each having a complex Gaussian distribution, $\mathbb{C}N(0,1)$

Each entry has probability density function,

$$f(u) = \pi^{-1} e^{-|u|^2}, \quad u \in \mathbb{C}$$

Shannon: Information theory

Channel capacity: “The tightest upper bound on the amount of information that can be transmitted reliably over a communications channel.”

Channel capacity is measured in “nats” (natural units for information entropy): 1 nat = $1/\log 2 \simeq 1.44$ bits

For MIMO models, the channel capacity is

$$I(\boldsymbol{x}, \boldsymbol{\varepsilon} | \boldsymbol{G}) = \log\det(\boldsymbol{I} + \boldsymbol{G}^* \boldsymbol{G})$$

where $\boldsymbol{G}^* = \overline{\boldsymbol{G}}'$ is the adjoint of $\boldsymbol{G}$

Tulino and Verdú, “Random Matrix Theory and Wireless Communications,” 2004

Average channel capacity: The average amount of information that can be transmitted by the channel.

Engineers wish to calculate the average channel capacity

We need to calculate $\mathbb{E}\, I(\boldsymbol{x}, \boldsymbol{\varepsilon} | \boldsymbol{G})$ with respect to the probability distribution of $\boldsymbol{G}$, $\boldsymbol{x}$, and $\boldsymbol{\varepsilon}$

Calculate the moment-generating function of $I(\boldsymbol{x}, \boldsymbol{\varepsilon} | \boldsymbol{G})$:

$$\mathbb{E}\left(e^{z\, I(\boldsymbol{x}, \boldsymbol{\varepsilon} | \boldsymbol{G})}\right), \qquad z \in \mathbb{C}$$

and then compute the first moment in the usual way:

$$\mathbb{E}\, I(\boldsymbol{x}, \boldsymbol{\varepsilon} | \boldsymbol{G}) = \frac{\partial}{\partial z} \mathbb{E}\left(e^{z\, I(\boldsymbol{x}, \boldsymbol{\varepsilon} | \boldsymbol{G})}\right)\Big|_{z=0}$$

Recall: $I(\boldsymbol{x}, \boldsymbol{\varepsilon} | \boldsymbol{G}) = \log\det(\boldsymbol{I} + \boldsymbol{G}^*\boldsymbol{G})$

$$\mathbb{E}\,(e^{z\,I(\boldsymbol{x},\boldsymbol{\varepsilon}|\boldsymbol{G})}) = \mathbb{E}\det(\boldsymbol{I} + \boldsymbol{G}^*\boldsymbol{G})^z$$

$\boldsymbol{G}$ has i.i.d. $\mathbb{C}N(0,1)$-distributed components, so

$$\mathbb{E}\det(\boldsymbol{I} + \boldsymbol{G}^*\boldsymbol{G})^z \propto \int_{\mathbb{C}^{n_r \times n_t}} \det(\boldsymbol{I} + \boldsymbol{G}^*\boldsymbol{G})^z\, e^{-\operatorname{tr}\boldsymbol{G}^*\boldsymbol{G}}\,\mathrm{d}\boldsymbol{G}$$

$\boldsymbol{G}^*\boldsymbol{G}$ has a complex Wishart distribution, so we can transform the integral into a Selberg integral over the eigenvalues of $\boldsymbol{G}^*\boldsymbol{G}$

$$\mathbb{E}\det(\boldsymbol{I} + \boldsymbol{G}^*\boldsymbol{G})^z \propto \int_{\mathbb{R}_+^{n_t}} \prod_{j<k}(\lambda_j - \lambda_k)^2 \prod_{k=1}^{n_t}(1+\lambda_k)^z\,\lambda_k^{n_t - n_r}\,e^{-\lambda_k}\,\mathrm{d}\lambda_k$$

Or, by changing variables to a Hermitian positive definite matrix $\boldsymbol{H} = \boldsymbol{G}^*\boldsymbol{G}$, we obtain

$$\mathbb{E}\,\det(\boldsymbol{I} + \boldsymbol{G}^*\boldsymbol{G})^z \propto \int_{\boldsymbol{H}>0} \det(\boldsymbol{I} + \boldsymbol{H})^z \,\det(\boldsymbol{H})^{n_t - n_r}\, e^{-\operatorname{tr}\boldsymbol{H}}\, \mathrm{d}\boldsymbol{H}$$

This integral is a confluent hypergeometric function (of the second kind) of matrix argument:

$$\mathbb{E}\,\det(\boldsymbol{I} + \boldsymbol{G}^*\boldsymbol{G})^z \propto \Psi(n_t; z + n_t; \boldsymbol{I})$$

J. Approx. Theory 1989 paper: The hypergeometric functions of Hermitian matrix argument expressed as ratios of determinants of classical hypergeometric functions.

The ratios-of-determinant formulas had been derived earlier by C. G. Khatri.

Using these results, the wireless communications community finally were able to obtain exact formulas for average channel capacity.

Let us now consider channels with more general probability distributions

Semi-correlated channels: Non-zero correlation among the transmitting antennas

The channel matrix $\boldsymbol{G}$ has probability density function

$$f(\boldsymbol{G}) \propto e^{-\operatorname{tr} \boldsymbol{\Sigma}^{-1} \boldsymbol{G}^* \boldsymbol{G}},$$

where $\boldsymbol{\Sigma}$ is a fixed Hermitian positive definite matrix.

$\boldsymbol{G}^* \boldsymbol{G}$ has a Wishart distribution with covariance matrix $\boldsymbol{\Sigma}$

The normalizing constant for $f(\boldsymbol{G})$ is $\det(\boldsymbol{\Sigma})^{-n_t} / \Gamma_{n_r}(n_t)$, where $\Gamma_{n_r}(n_t)$ is a product of classical Gamma functions.

The moment-generating function of channel capacity:

$$\mathbb{E} \det(\boldsymbol{I} + \boldsymbol{G}^* \boldsymbol{G})^z \propto \Psi(n_t; z + n_t; \boldsymbol{\Sigma})$$

Fully correlated channels: Non-zero correlation among the transmitting and the receiving antennas

The channel matrix $\boldsymbol{G}$ has the probability density function

$$f(\boldsymbol{G}) \propto e^{-\mathrm{tr}\,\boldsymbol{\Sigma}^{-1}\boldsymbol{G}^*\boldsymbol{\Lambda}^{-1}\boldsymbol{G}},$$

where $\boldsymbol{\Sigma}$ and $\boldsymbol{\Lambda}$ are fixed Hermitian positive definite matrices.

The LDU decomposition of $\boldsymbol{G}$ leads to integrals over unitary groups

The moment-generating function of $I(\boldsymbol{G})$ is obtained in terms of a hypergeometric function of two matrix arguments

The formulas in the 1989 J.A.T. paper again lead to exact calculation of the average channel capacity

Non-centered channels

In this case, $\mathbb{E}(\boldsymbol{G}) = \boldsymbol{G}_0$, a non-zero matrix

The channel matrix $\boldsymbol{G}$ has probability density function

$$f(\boldsymbol{G}) \propto e^{-\mathrm{tr}\,(\boldsymbol{G}-\boldsymbol{G}_0)^*(\boldsymbol{G}-\boldsymbol{G}_0)}$$

$\boldsymbol{G}^*\boldsymbol{G}$ has a non-central Wishart distribution, and the normalizing constant involves a Bessel function of matrix argument

Herz (1955)

$U(n)$: The group of unitary $n \times n$ matrices

$\boldsymbol{U}$: A generic element of $U(n)$

$\mathrm{d}\boldsymbol{U}$: Haar measure on $U(n)$, normalized to have total volume 1

All of the matrix integrals that arise here stem from the integral

$$\int_{U(n)} e^{\mathrm{tr}\, \boldsymbol{S}\boldsymbol{U}\boldsymbol{T}\boldsymbol{U}^*} \, \mathrm{d}\boldsymbol{U}$$

where $\boldsymbol{S}$ and $\boldsymbol{T}$ are Hermitian $n \times n$ matrices

How do we calculate this integral?

Partition: $\mu = (\mu_1, \ldots, \mu_n)$, integers, where $\mu_1 \geq \cdots \geq \mu_n \geq 0$

Length of μ: The number of non-zero μ_j

Weight of μ: $|\mu| = \mu_1 + \cdots + \mu_n$

Schur function: For $t_1, \ldots, t_n \in \mathbb{C}$,

$$\chi_\mu(t_1, \ldots, t_n) = \frac{\det({t_i}^{\mu_j+n-j})}{V(t_1, \ldots, t_n)}$$

where

$$V(t_1, \ldots, t_n) = \prod_{1 \leq i < j \leq n} (t_i - t_j)$$

$\{\chi_\mu(t_1, \ldots, t_n) : |\mu| = j\}$ is a basis for the space of polynomials in $t_1, \ldots, t_n$ which are symmetric and homogeneous of degree j

χ_μ satisfies a remarkable number of properties:

Characters of irreducible representations of $U(n)$

Spherical functions on the cone of Hermitian p.d. matrices

$$d_\mu := \chi_\mu(1,\ldots,1) = \lim_{t_1,\ldots,t_n \to 1} \frac{\det({t_i}^{\mu_j+n-j})}{V(t_1,\ldots,t_n)}$$

$$d_\mu = \beta_n^{-1} \prod_{1\le i<j\le n} (\mu_i - \mu_j - i + j)$$

where $\beta_n = 0!1!2!\cdots(n-1)!$

Extend χ_μ to the space of Hermitian matrices

$$\chi_\mu(\boldsymbol{T}) = \chi_\mu(t_1, \ldots, t_n)$$

where $t_1, \ldots, t_n$ are the eigenvalues of $\boldsymbol{T}$

$\chi_\mu(\boldsymbol{T})$ is homogeneous and unitarily invariant:

$$\chi_\mu(\boldsymbol{UTU}^*) = \chi_\mu(\boldsymbol{T})$$

for all unitary $n \times n$ matrices $\boldsymbol{U}$

There are constants ω_μ such that

$$(\mathrm{tr}\,\boldsymbol{T})^j = \sum_{|\mu|=j} \omega_\mu \chi_\mu(\boldsymbol{T})$$

$$\omega_\mu = |\mu|!\,\frac{\prod_{1\le i<j\le n} (\mu_i - \mu_j - i + j)}{\prod_{j=1}^{n} (\mu_j + n - j)!}$$

Define $Z_\mu(\boldsymbol{T}) = \omega_\mu \chi_\mu(\boldsymbol{T})$: The (complex) zonal polynomial

$$(\mathrm{tr}\,\boldsymbol{T})^j = \sum_{|\mu|=j} Z_\mu(\boldsymbol{T})$$

$$e^{\mathrm{tr}\,\boldsymbol{T}} = \sum_{j=0}^{\infty} \frac{1}{j!} \sum_{|\mu|=j} Z_\mu(\boldsymbol{T})$$

$$e^{\operatorname{tr} \boldsymbol{SUTU}^*} = \sum_{j=0}^{\infty} \frac{1}{j!} \sum_{|\mu|=j} Z_\mu(\boldsymbol{SUTU}^*)$$

Because χ_μ is a spherical function,

$$\int_{U(n)} Z_\mu(\boldsymbol{SUTU}^*)\, \mathrm{d}\boldsymbol{U} = \frac{Z_\mu(\boldsymbol{S})\, Z_\mu(\boldsymbol{T})}{Z_\mu(\boldsymbol{I})}$$

$$\int_{U(n)} e^{\operatorname{tr} \boldsymbol{SUTU}^*}\, \mathrm{d}\boldsymbol{U} = \sum_{j=0}^{\infty} \frac{1}{j!} \sum_{|\mu|=j} \frac{Z_\mu(\boldsymbol{S})\, Z_\mu(\boldsymbol{T})}{Z_\mu(\boldsymbol{I})}$$

Apply the explicit formulas for Z_μ, ω_μ, and d_μ

$$\int_{U(n)} e^{\operatorname{tr} \boldsymbol{S U T U}^*} \, \mathrm{d}\boldsymbol{U}$$

$$= \beta_n^{-1} \sum_{\mu} \frac{\det(s_i^{\mu_j+n-j})}{V(\boldsymbol{S})} \frac{\det(t_i^{\mu_j+n-j})}{V(\boldsymbol{T})} \prod_{j=1}^{n} \frac{1}{(\mu_j+n-j)!}$$

Set $k_j = \mu_j + n - j$, $j = 1, \ldots, n$

Then $k_1 > k_2 > \cdots > k_n \geq 0$, so the sum over μ reduces to

$$\sum_{k_1 > \cdots > k_n \geq 0} \det(s_i^{k_j}) \det(t_i^{k_j}) \prod_{j=1}^{n} \frac{1}{k_j!}$$

A weighted sum of the product of determinants

The Binet-Cauchy formula

$$\det\left(\int f_i(x) g_j(x)\, \mathrm{d}\nu(x)\right) = \int_{x_1 > \cdots > x_n} \cdots \int \det(f_i(x_j))\, \det(g_i(x_j)) \prod_{j=1}^{n} \mathrm{d}\nu(x_j)$$

Apply the Binet-Cauchy formula with ν as a discrete measure on the nonnegative integers to obtain the explicit formula

$$\int_{U(n)} e^{\mathrm{tr}\, \boldsymbol{S}\boldsymbol{U}\boldsymbol{T}\boldsymbol{U}^*}\, \mathrm{d}\boldsymbol{U} = \beta_n^{-1} \frac{\det\left(e^{s_i t_j}\right)}{V(\boldsymbol{S})\, V(\boldsymbol{T})}$$

Remarks on the representation theory of $U(n)$ and related explicit formulas for all hypergeometric functions of matrix argument