Operator-Valued Free Probability and Block Random Matrices

Roland Speicher

Saarland University
Saarbrücken, Germany

Asymptotic Freeness of Random Matrices

Basic Observation (Voiculescu 1991)

Large classes of independent random matrices (like Gaussian or Wishart matrices) become asymptoticially freely independent.

Section 1

Free Probability Theory

Definition (Voiculescu 1985)

Let $(\mathcal{A}, \varphi)$ be a **non-commutative probability space**, i.e., $\mathcal{A}$ is a unital algebra and $\varphi : \mathcal{A} \to \mathbb{C}$ is unital linear functional (i.e., $\varphi(1) = 1$).

Example (Commutative Probability Space)

For a classical probability space (Ω, P) take

- $\mathcal{A} = L^\infty(\Omega, P)$
- $\varphi(x) = \int_\Omega x(\omega) dP(\omega)$ for $x \in \mathcal{A}$

Definition (Voiculescu 1985)

Let $(\mathcal{A}, \varphi)$ be a **non-commutative probability space**, i.e., $\mathcal{A}$ is a unital algebra and $\varphi : \mathcal{A} \to \mathbb{C}$ is unital linear functional (i.e., $\varphi(1) = 1$).

Unital subalgebras $\mathcal{A}_i$ $(i \in I)$ are **free** or **freely independent**, if $\varphi(a_1 \cdots a_n) = 0$ whenever

- $a_i \in \mathcal{A}_{j(i)} \qquad j(i) \in I \quad \forall i$
- $j(1) \neq j(2) \neq \cdots \neq j(n)$
- $\varphi(a_i) = 0 \quad \forall i$

Random variables $x_1, \dots, x_n \in \mathcal{A}$ are freely independent, if their generated unital subalgebras $\mathcal{A}_i :=$ algebra$(1, x_i)$ are so.

What is Freeness?

Freeness between x and y is an infinite set of equations relating various moments in x and y:

$$\varphi\Big(p_1(x)q_1(y)p_2(x)q_2(y)\cdots\Big) = 0$$

Basic observation: free independence between x and y is actually a **rule for calculating mixed moments** in x and y from the moments of x and the moments of y:

$$\varphi\Big(x^{m_1}y^{n_1}x^{m_2}y^{n_2}\cdots\Big) = \text{polynomial}\big(\varphi(x^i), \varphi(y^j)\big)$$

Example

If x and y are freely independent, then we have

$$\varphi(x^m y^n) = \varphi(x^m)\cdot\varphi(y^n)$$

Example

If x and y are freely independent, then we have

$$\varphi(x^m y^n) = \varphi(x^m) \cdot \varphi(y^n)$$

$$\varphi(x^{m_1} y^n x^{m_2}) = \varphi(x^{m_1+m_2}) \cdot \varphi(y^n)$$

but also

$$\varphi(xyxy) = \varphi(x^2) \cdot \varphi(y)^2 + \varphi(x)^2 \cdot \varphi(y^2) - \varphi(x)^2 \cdot \varphi(y)^2$$

Free independence is a rule for calculating mixed moments, analogous to the concept of independence for random variables.
Note: free independence is a different rule from classical independence; free independence occurs typically for **non-commuting random variables**, like operators on Hilbert spaces.

Example

If x and y are freely independent, then we have

$$\varphi(x^m y^n) = \varphi(x^m) \cdot \varphi(y^n)$$

$$\varphi(x^{m_1} y^n x^{m_2}) = \varphi(x^{m_1+m_2}) \cdot \varphi(y^n)$$

but also

$$\varphi(xyxy) = \varphi(x^2) \cdot \varphi(y)^2 + \varphi(x)^2 \cdot \varphi(y^2) - \varphi(x)^2 \cdot \varphi(y)^2$$

This means of course that, for any polynomial p, the moments of $p(x,y)$ are determined in terms of the moments of x and the moments of y.

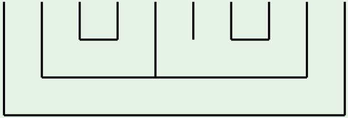
Combinatorial Structure of Freeness

Basic Observation (Speicher 1993)

The structure of the formulas for mixed moments is governed by the **lattice of non-crossing partitions**.

Example (Factorization of Non-Crossing Moments)

Let $x_1, \dots, x_5$ be free. Consider $x_1\, x_2\, x_3\, x_3\, x_2\, x_4\, x_5\, x_5\, x_2 x_1$ Then



$$\varphi(x_1 x_2 x_3 x_3 x_2 x_4 x_5 x_5 x_2 x_1)$$
$$= \varphi(x_1 x_1) \cdot \varphi(x_2 x_2 x_2) \cdot \varphi(x_3 x_3) \cdot \varphi(x_4) \cdot \varphi(x_5 x_5)$$

Crossing moments are more complicated, but still have non-crossing structure: $\varphi(xyxy) = \varphi(x^2) \cdot \varphi(y)^2 + \varphi(x)^2 \cdot \varphi(y^2) - \varphi(x)^2 \cdot \varphi(y)^2$

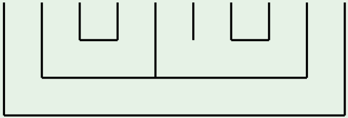
Combinatorial Structure of Freeness

Basic Observation (Speicher 1993)

The structure of the formulas for mixed moments is governed by the **lattice of non-crossing partitions**.

Example (Factorization of Non-Crossing Moments)

Let $x_1, \dots, x_5$ be free. Consider $x_1\, x_2\, x_3\, x_3\, x_2\, x_4\, x_5\, x_5\, x_2 x_1$ Then



$$\varphi(x_1 x_2 x_3 x_3 x_2 x_4 x_5 x_5 x_2 x_1)$$
$$= \varphi(x_1 x_1) \cdot \varphi(x_2 x_2 x_2) \cdot \varphi(x_3 x_3) \cdot \varphi(x_4) \cdot \varphi(x_5 x_5)$$

- Many consequences of this are worked out in joint works with A. Nica
- Nica, Speicher: *Lectures on the Combinatorics of Free Probability*, 2006

Where Does Free Independence Show Up?

Free independence can be found in different situations; some of the main occurrences are:

- generators of the free group in the corresponding free group von Neumann algebras $L(\mathbb{F}_n)$
- creation and annihilation operators on full Fock spaces
- for many classes of random matrices

Asymptotic Freeness of Random Matrices

Theorem (Voiculescu 1991)

Large classes of independent random matrices (like Gaussian or Wishart matrices) become asymptoticially freely independent, with respect to $\varphi = tr := \frac{1}{N} Tr$, if $N \to \infty$.

Example

This means, for example: if X_N and Y_N are independent $N \times N$ Wigner and Wishart matrices, respectively, then we have almost surely:

$$\lim_{N\to\infty} \text{tr}(X_N Y_N X_N Y_N) = \lim_{N\to\infty} \text{tr}(X_N^2) \cdot \lim_{N\to\infty} \text{tr}(Y_N)^2 + \lim_{N\to\infty} \text{tr}(X_N)^2 \cdot \lim_{N\to\infty} \text{tr}(Y_N^2) - \lim_{N\to\infty} \text{tr}(X_N)^2 \cdot \lim_{N\to\infty} \text{tr}(Y_N)^2$$

Hence we have a rule for calculating asymptotically mixed moments of our matrices with respect to the normalized trace tr.

Asymptotic Freeness of Random Matrices

Theorem (Voiculescu 1991)

Large classes of independent random matrices (like Gaussian or Wishart matrices) become asymptoticially freely independent, with respect to $\varphi = tr := \frac{1}{N} Tr$, if $N \to \infty$.

Note that moments with respect to tr determine the eigenvalue distribution of a matrix.
For an $N \times N$ matrix $X = X^*$ with eigenvalues $\lambda_1, \dots, \lambda_N$ its eigenvalue distribution

$$\mu_X := \frac{1}{N}(\delta_{\lambda_1} + \cdots + \delta_{\lambda_N})$$

is determined by

$$\int_{\mathbb{R}} t^k d\mu_X(t) = \mathrm{tr}(X^k) \qquad \text{for all } k = 0, 1, 2, \dots$$

.

Section 2

Free Convolution

Sum of Free Variables

Consider x, y free.
Then, by freeness, the moments of $x + y$ are uniquely determined by the moments of x and the moments of y.

Notation

We say the distribution of $x + y$ is the

free convolution

of the distribution of x and the distribution of y,

$$\mu_{x+y} = \mu_x \boxplus \mu_y.$$

The Cauchy Transform

Definition

For any probability measure μ on $\mathbb{R}$ we define its **Cauchy transform** by

$$G(z) := \int_{\mathbb{R}} \frac{1}{z-t} d\mu(t)$$

$-G$ is also called **Stieltjes transform**.

This is an analytic function $G : \mathbb{C}^+ \to \mathbb{C}^-$ and we can recover μ from G by **Stieltjes inversion formula**.

$$d\mu(t) = -\frac{1}{\pi} \lim_{\varepsilon \to 0} \Im G(t + i\varepsilon) dt$$

The R-transform

Definition

Consider a random variable $x \in \mathcal{A}$. Let G be its Cauchy transform

$$G(z) = \varphi[\frac{1}{z-x}] = \frac{1}{z} + \sum_{n=1}^{\infty} \frac{\varphi(x^n)}{z^{n+1}}.$$

We define its ***R*-transform** by the equation

$$\frac{1}{G(z)} + R[G(z)] = z$$

Theorem (Voiculescu 1986)

The R-transform linearizes free convolution, i.e.,

$$R_{x+y}(z) = R_x(z) + R_y(z) \qquad \text{if } x \text{ and } y \text{ are free.}$$

Calculation of Free Convolution by R-transform

The relation between Cauchy transform and R-transform, and the Stieltjes inversion formula give an effective algorithm for calculating free convolutions; and thus also, e.g., the asymptotic eigenvalue distribution of sums of random matrices in generic position:

$$
\begin{array}{ccccccccc}
x & \rightsquigarrow & G_x & \rightsquigarrow & R_x & & & & \\
 & & & & \downarrow & & & & \\
 & & & & R_x + R_y = R_{x+y} & \rightsquigarrow & G_{x+y} & \rightsquigarrow & x+y \\
 & & & & \uparrow & & & & \\
y & \rightsquigarrow & G_y & \rightsquigarrow & R_y & & & &
\end{array}
$$

What is the Free Binomial $\left(\frac{1}{2}\delta_{-1}+\frac{1}{2}\delta_{+1}\right)^{\boxplus 2}$

Example

$$\mu := \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_{+1}, \qquad \nu := \mu \boxplus \mu$$

Then $$G_\mu(z) = \int \frac{1}{z-t} d\mu(t) = \frac{1}{2}\Big(\frac{1}{z+1} + \frac{1}{z-1}\Big) = \frac{z}{z^2-1}$$

and so $$z = G_\mu[R_\mu(z)+1/z] = \frac{R_\mu(z)+1/z}{(R_\mu(z)+1/z)^2-1}$$

thus $$R_\mu(z) = \frac{\sqrt{1+4z^2}-1}{2z}$$

and so $$R_\nu(z) = 2R_\mu(z) = \frac{\sqrt{1+4z^2}-1}{z}$$

What is the Free Binomial $\left(\frac{1}{2}\delta_{-1}+\frac{1}{2}\delta_{+1}\right)^{\boxplus 2}$

Example

$$R_\nu(z)=\frac{\sqrt{1+4z^2}-1}{z} \qquad \text{gives} \qquad G_\nu(z)=\frac{1}{\sqrt{z^2-4}}$$

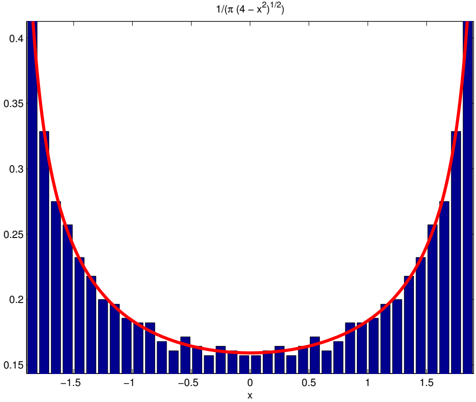
and thus

$$d\nu(t)=-\frac{1}{\pi}\Im\frac{1}{\sqrt{t^2-4}}dt=\begin{cases}\frac{1}{\pi\sqrt{4-t^2}}, & |t|\le 2\\ 0, & \text{otherwise}\end{cases}$$

So

$$\left(\frac{1}{2}\delta_{-1}+\frac{1}{2}\delta_{+1}\right)^{\boxplus 2}=\nu=\text{arcsine-distribution}$$

What is the Free Binomial $\left(\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_{+1}\right)^{\boxplus 2}$



2800 eigenvalues of $A + UBU^*$, where A and B are diagonal matrices with 1400 eigenvalues +1 and 1400 eigenvalues -1, and U is a randomly chosen unitary matrix

The R-transform as an Analytic Object

- The R-transform can be established as an analytic function via power series expansions around the point infinity in the complex plane.
- The R-transform can, in contrast to the Cauchy transform, in general not be defined on all of the upper complex half-plane, but only in some truncated cones (which depend on the considered variable).
- The equation $\frac{1}{G(z)} + R[G(z)] = z$ does in general not allow explicit solutions and there is no good numerical algorithm for dealing with this.

Problem

The R-transform is not really an adequate analytic tool for more complicated problems.
Is there an alternative?

An Alternative to the R-transform: Subordination

Let x and y be free. Put $w := R_{x+y}(z) + 1/z$, then

$$G_{x+y}(w) = z = G_x[R^x(z)+1/z] = G_x[w-R_y(z)] = G_x[w-R_y[G_{x+y}(w)]]$$

Basic Observation (Voiculescu, Biane, Götze, Chistyakov, Belinschi, Bercovici ...)

There are nice analytic descriptions in **subordination form**, e.g., for x and y free one has

$$G_{x+y}(z) = G_x\big(\omega(z)\big),$$

where $\omega : \mathbb{C}^+ \to \mathbb{C}^+$ is an analytic function which can be calculated effectively via fixpoint descriptions.

The Subordination Function

Theorem (Belinschi, Bercovici 2007)

Let x and y be free. Put

$$F(z) := \frac{1}{G(z)}$$

Then there exists an analytic $\omega : \mathbb{C}^+ \to \mathbb{C}^+$ such that

$$F_{x+y}(z) = F_x\big(\omega(z)\big) \qquad \text{and} \qquad G_{x+y}(z) = G_x\big(\omega(z)\big)$$

The subordination function $\omega(z)$ is given as the unique fixed point in the upper half-plane of the map

$$f_z(w) = F_y(F_x(w) - w + z) - (F_x(w) - w)$$

Example: semicircle ⊞ Marchenko-Pastur

Example

Let s be semicircle, p be Marchenko-Pastur (i.e., free Poisson) and s, p free. Consider $a := s + p$.

$$R_s(z) = z, \quad R_p(z) = \frac{\lambda}{1-z},$$

thus we have

$$R_a(z) = R_s(z) + R_p(z) = z + \frac{\lambda}{1-z},$$

and hence

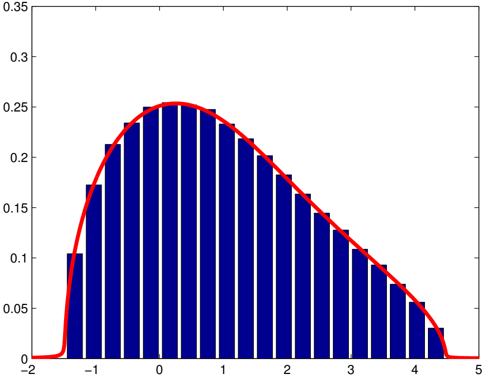
$$G_a(z) + \frac{\lambda}{1 - G_a(z)} + \frac{1}{G_a(z)} = z$$

Alternative subordination formulation

$$G_{s+p}(z) = G_p\big[z - R_s[G_{s+p}(z)]\big] = G_p[z - G_{s+p}(z)]$$

Example: semicircle ⊞ Marchenko-Pastur

$$G_{s+p}(z) = G_p\big[z - R_s[G_{s+p}(z)]\big] = G_p[z - G_{s+p}(z)]$$



Section 3

Gaussian Random Matrices and Semicircular Element

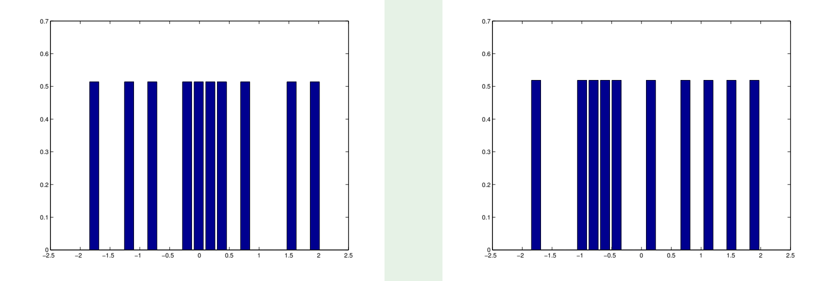
Gaussian Random Matrix (Wigner 1955)

Definition

A **Gaussian random matrix** $A_N = \frac{1}{\sqrt{N}}\left(x_{ij}\right)_{i,j=1}^{N}$

- is symmetric: $A_N^* = A_N$
- $\{x_{ij} \mid 1 \leq i \leq j \leq N\}$ are independent and identically distributed, with a centered normal distribution of variance 1

Example (eigenvalue distribution for $N = 10$)



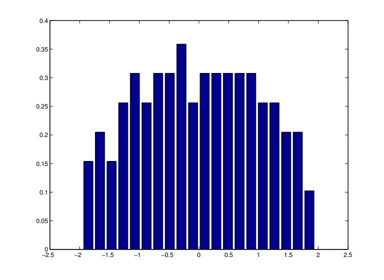
Gaussian Random Matrix (Wigner 1955)

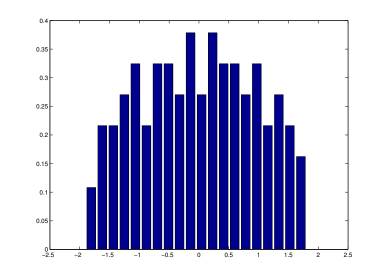
Definition

A **Gaussian random matrix** $A_N = \frac{1}{\sqrt{N}}(x_{ij})_{i,j=1}^{N}$

- is symmetric: $A_N^* = A_N$
- $\{x_{ij} \mid 1 \leq i \leq j \leq N\}$ are independent and identically distributed, with a centered normal distribution of variance 1

Example (eigenvalue distribution for $N = 100$)





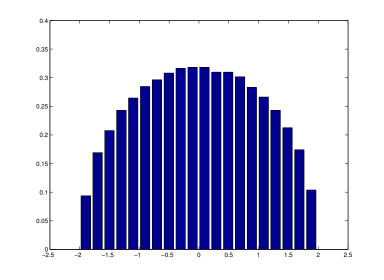
Gaussian Random Matrix (Wigner 1955)

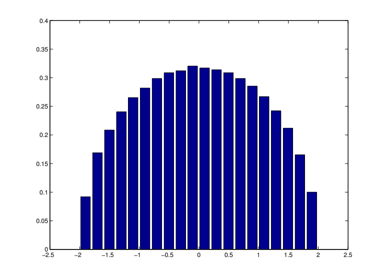
Definition

A **Gaussian random matrix** $A_N = \frac{1}{\sqrt{N}}\left(x_{ij}\right)_{i,j=1}^{N}$

- is symmetric: $A_N^* = A_N$
- $\{x_{ij} \mid 1 \leq i \leq j \leq N\}$ are independent and identically distributed, with a centered normal distribution of variance 1

Example (eigenvalue distribution for $N = 3000$)





Definition

The **empirical eigenvalue distribution** of A_N is

$$\mu_{A_N}(\omega) = \frac{1}{N} \sum_{i=1}^{N} \delta_{\lambda_i(\omega)}$$

where $\lambda_i(\omega)$ are the N eigenvalues (counted with multiplicity) of $A_N(\omega)$

Theorem (Wigner's semicircle law)

We have almost surely

$$\boldsymbol{\mu_{A_N} \Longrightarrow \mu_W} \qquad \textit{(weak convergence)}$$

i.e., for each continuous and bounded f we have almost surely

$$\lim_{N\to\infty} \int_{\mathbb{R}} f(t) d\mu_{A_N}(t) = \int_{\mathbb{R}} f(t) d\mu_W(t) = \frac{1}{2\pi} \int_{-2}^{2} f(t) \sqrt{4 - t^2} dt$$

Proof of the Semicircle Law

One shows

$$\lim_{N\to\infty} \mu_{A_N}(f) = \mu_W(f) \qquad \text{almost surely}$$

in two steps:

- convergence in average:

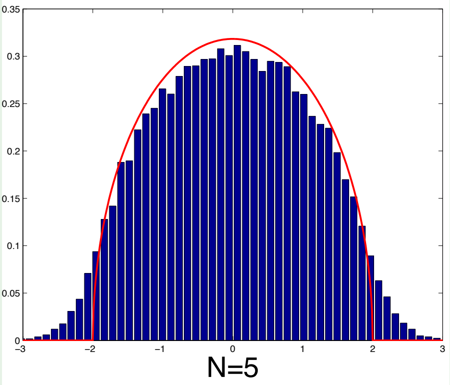
$$\lim_{N\to\infty} E[\mu_{A_N}(f)] = \mu_W(f)$$

- fluctuations are negligible for $N \to \infty$:

$$\sum_N \mathrm{Var}[\mu_{A_N}(f)] < \infty$$

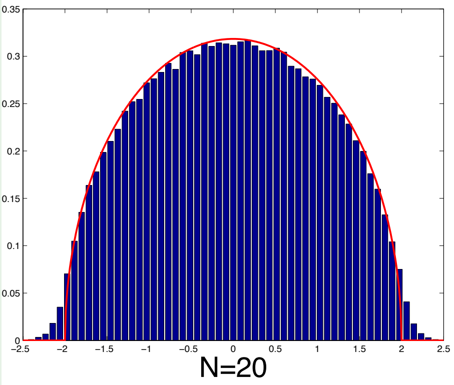
Convergence of Averaged Eigenvalue Distribution

Example (eigenvalue distribution for $N = 5$)



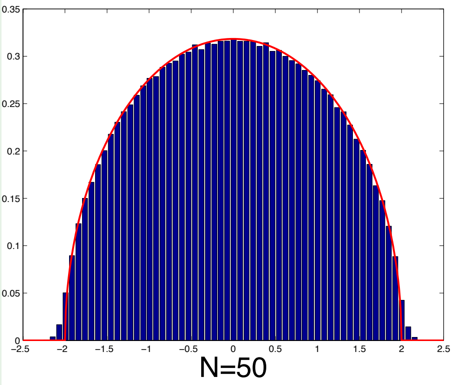
Convergence of Averaged Eigenvalue Distribution

Example (eigenvalue distribution for $N = 20$)



Convergence of Averaged Eigenvalue Distribution

Example (eigenvalue distribution for $N = 50$)



Convergence in Average

For

$$\lim_{N\to\infty} E[\mu_{A_N}(f)] = \mu_W(f)$$

it suffices to treat **convergence of all averaged moments**, i.e.,

$$\lim_{N\to\infty} E[\int t^n d\mu_{A_N}(t)] = \int t^n d\mu_W(t) \qquad \forall n \in \mathbb{N}$$

Note:

$$E[\int t^n d\mu_{A_N}(t)] = E[\frac{1}{N}\sum_{i=1}^{N} \lambda_i^n] = E[\text{tr}(A_N^n)]$$

Calculation of Averaged Moments

Note:

$$E[\int t^n d\mu_{A_N}(t)] = E[\frac{1}{N}\sum_{i=1}^{N}\lambda_i^n] = E[\mathrm{tr}(A_N^n)]$$

but

$$E[\mathrm{tr}(A_N^n)] = \frac{1}{N}\sum_{i_1,\dots,i_n=1}^{N} E[a_{i_1 i_2} a_{i_2 i_3} \cdots a_{i_n i_1}]$$

Calculation of Averaged Moments

Note:

$$E[\int t^n d\mu_{A_N}(t)] = E[\frac{1}{N}\sum_{i=1}^{N}\lambda_i^n] = E[\mathrm{tr}(A_N^n)]$$

but

$$E[\mathrm{tr}(A_N^n)] = \frac{1}{N}\sum_{i_1,\dots,i_n=1}^{N} \underbrace{E[a_{i_1 i_2} a_{i_2 i_3} \cdots a_{i_n i_1}]}_{\substack{\text{expressed in} \\ \text{terms of pairings} \\ \text{``Wick formula''}}}$$

Semicircular Element

Asymptotically, for $N \to \infty$, only **non-crossing pairings** survive:

$$\lim_{N\to\infty} E[\mathrm{tr}(A_N^n)] = \#NC_2(n)$$

Definition

Define limiting **semicircle element** s by

$$\varphi(s^n) := \#NC_2(n).$$

($s \in \mathcal{A}$, where $\mathcal{A}$ is some unital algebra, $\varphi : \mathcal{A} \to \mathbb{C}$)

Notation

Then we say that our Gaussian random matrices A_N converge in distribution to the semicircle element s,

$$A_N \xrightarrow{\text{distr}} s$$

What is Distribution of s?

$$\varphi(s^n) = \lim_{N\to\infty} E[\text{tr}(A_N^n)] = \#NC_2(n)$$

Claim

$$\varphi(s^n) = \int t^n d\mu_W(t)$$

more concretely:

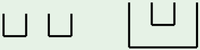
$$\#NC_2(n) = \frac{1}{2\pi}\int_{-2}^{+2} t^n \sqrt{4-t^2}\,dt$$

What is Distribution of s?

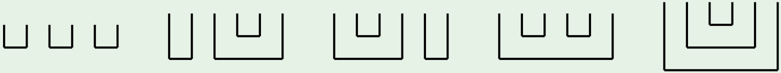
Example

$n = 2$: $\varphi(s^2) = 1$

$n = 4$: $\varphi(s^4) = 2$



$n = 6$: $\varphi(s^6) = 5$



What is Distribution of s?

Claim

$$\varphi(s^{2k}) = C_k \qquad k\text{-th Catalan number}$$

What are the **Catalan numbers**?

- $C_k = \frac{1}{k+1}\binom{2k}{k}$
- C_k is determined by $C_0 = C_1 = 1$ and the recurrence relation

$$C_k = \sum_{l=1}^{k} C_{l-1}C_{k-l}.$$

Moments of s are Given by Catalan Numbers

It is fairly easy to see that the moments $\varphi(s^{2k})$ satisfy the recursion for the Catalan numbers:

$$\varphi(s^{2k}) = \sum_{l=1}^{k} \varphi(s^{2l-2})\varphi(s^{2k-2l}).$$

Notation

$$M(z) := \sum_{n=0}^{\infty} \varphi(s^n) z^n = 1 + \sum_{k=1}^{\infty} \varphi(s^{2k}) z^{2k}$$

$$\begin{aligned} M(z) &= 1 + z^2 \sum_{k=1}^{\infty} \sum_{l=1}^{k} \varphi(s^{2l-2}) z^{2l-2} \varphi(s^{2k-2l}) z^{2k-2l} \\ &= 1 + z^2 M(z) \cdot M(z) \end{aligned}$$

Moments of s are Given by Catalan Numbers

$$M(z) = 1 + z^2 M(z) \cdot M(z)$$

Notation (Cauchy transform)

Instead of moment generating series $M(z)$ consider

$$G(z) := \varphi\left(\frac{1}{z-s}\right)$$

Note

$$G(z) = \sum_{n=0}^{\infty} \frac{\varphi(s^n)}{z^{n+1}} = \frac{1}{z} \sum_{n=0}^{\infty} \varphi(s^n)\left(\frac{1}{z}\right)^n = \frac{1}{z} M(1/z),$$

thus

$$zG(z) = 1 + G(z)^2$$

For the basic Gaussian random matrix ensemble one can thus derive equations for the Cauchy transform of the limiting eigenvalue distribution, solve those equations and then get the density via Stieltjes inversion.

Example (Gaussian rm)

$$G(z)^2 + 1 = zG(z),$$

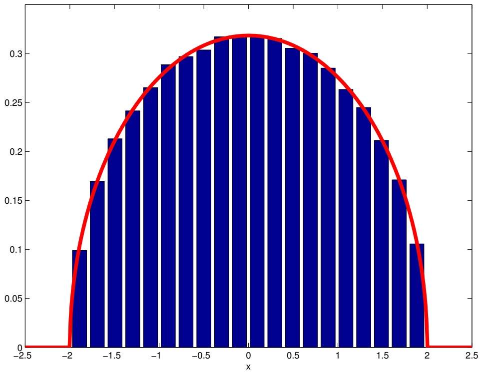
which can be solved as

$$G(z) = \frac{z - \sqrt{z^2 - 4}}{2},$$

thus

$$d\mu_s(t) = \frac{1}{2\pi}\sqrt{4 - t^2}dt$$

Wigners semicircle



Section 4

Block Random Matrices and Operator-Valued Semicircular Elements

Eigenvalue Distribution of Block Matrices

Example

Consider the **block matrix**

$$X_N = \begin{pmatrix} A_N & B_N & C_N \\ B_N & A_N & B_N \\ C_N & B_N & A_N \end{pmatrix},$$

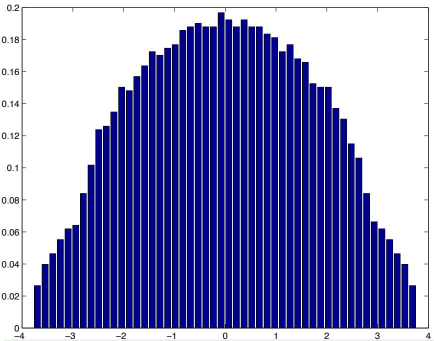
where A_N, B_N, C_N are independent Gaussian $N \times N$-random matrices.

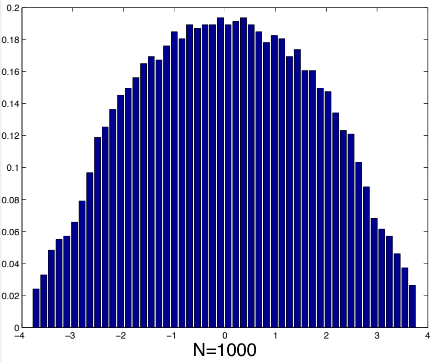
Problem

What is eigenvalue distribution of X_N for $N \to \infty$?

Typical Eigenvalue Distribution for $N = 1000$

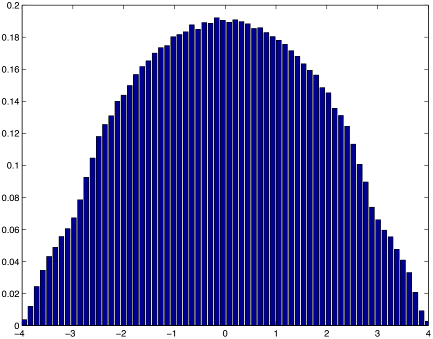
Example



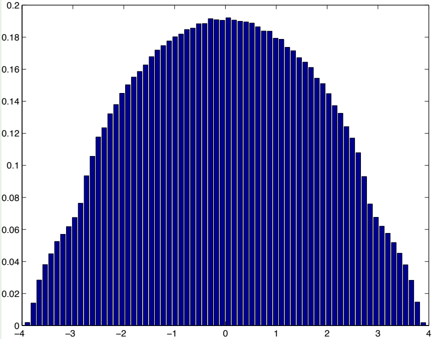


Averaged Eigenvalue Distribution

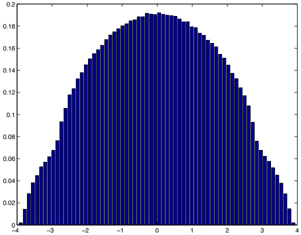
Example



N=50



N=150



Problem

This limiting distribution is not a semicircle, and it cannot be described nicely within usual free probability theory.

Solution

However, it fits well into the frame of

operator-valued free probability theory!

What is an operator-valued probability space?

scalars	$\longrightarrow$	operator-valued scalars
$\mathbb{C}$		$\mathcal{B}$
state	$\longrightarrow$	conditional expectation
$\varphi : \mathcal{A} \to \mathbb{C}$		$E : \mathcal{A} \to \mathcal{B}$
		$E[b_1 a b_2] = b_1 E[a] b_2$
moments	$\longrightarrow$	operator-valued moments
$\varphi(a^n)$		$E[ab_1 ab_2 a \cdots ab_{n-1} a]$

Example: $M_2(\mathbb{C})$-valued probability space

Example

Let $(\mathcal{C}, \varphi)$ be a non-commutative probability space. Put

$$M_2(\mathcal{C}) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a, b, c, d \in \mathcal{C} \right\}$$

and consider $\psi := \text{tr} \otimes \varphi$ and $E := \text{id} \otimes \varphi$, i.e.:

$$\psi \left[\begin{pmatrix} a & b \\ c & d \end{pmatrix} \right] = \frac{1}{2}(\varphi(a) + \varphi(d)), \qquad E \left[\begin{pmatrix} a & b \\ c & d \end{pmatrix} \right] = \begin{pmatrix} \varphi(a) & \varphi(b) \\ \varphi(c) & \varphi(d) \end{pmatrix}$$

- $(M_2(\mathcal{C}), \psi)$ is a non-commutative probability space, and
- $(M_2(\mathcal{C}), E)$ is an $M_2(\mathbb{C})$-valued probability space

What is an operator-valued semicircular element?

Consider an operator-valued probability space

$$E : \mathcal{A} \to \mathcal{B}$$

Definition

$s \in \mathcal{A}$ is **semicircular** if

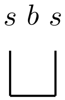
- second moment is given by

$$E[sbs] = \eta(b)$$

  for a completely positive map $\eta : \mathcal{B} \to \mathcal{B}$
- higher moments of s are given in terms of second moments by summing over non-crossing pairings

Moments of an Operator-Valued Semicircle

$$E[sbs] = \eta(b)$$



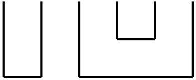
$$E[sb_1 sb_2 s \cdots sb_{n-1} s] = \sum_{\pi \in NC_2(n)} \Big(\text{iterated application of } \eta \text{ according to } \pi\Big)$$

Sixth Moment of Operator-Valued Semicircle

$sb_1sb_2sb_3sb_4sb_5s$

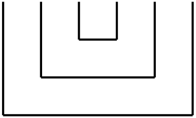
$\eta(b_1)\cdot b_2\cdot\eta(b_3)\cdot b_4\cdot\eta(b_5)$

$sb_1sb_2sb_3sb_4sb_5s$



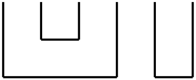
$\eta(b_1)\cdot b_2\cdot\eta\big(b_3\cdot\eta(b_4)\cdot b_5\big)$

$sb_1sb_2sb_3sb_4sb_5s$



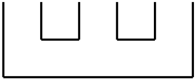
$\eta\Big(b_1\cdot\eta\big(b_2\cdot\eta(b_3)\cdot b_4\big)\cdot b_5\Big)$

$sb_1sb_2sb_3sb_4sb_5s$



$\eta\big(b_1\cdot\eta(b_2)\cdot b_3\big)\cdot b_4\cdot\eta(b_5)$

$sb_1sb_2sb_3sb_4sb_5s$



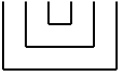
$\eta\big(b_1\cdot\eta(b_2)\cdot b_3\cdot\eta(b_4)\cdot b_5\big)$

Sixth Moment of Operator-Valued Semicircle

$$E[sb_1sb_2sb_3sb_4sb_5s] = \eta(b_1)\cdot b_2\cdot\eta(b_3)\cdot b_4\cdot\eta(b_5)$$

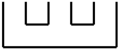
$$+\,\eta(b_1)\cdot b_2\cdot\eta\big(b_3\cdot\eta(b_4)\cdot b_5\big)$$

$$+\,\eta\Big(b_1\cdot\eta\big(b_2\cdot\eta(b_3)\cdot b_4\big)\cdot b_5\Big)$$



$$+\,\eta\big(b_1\cdot\eta(b_2)\cdot b_3\big)\cdot b_4\cdot\eta(b_5)$$

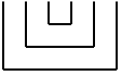
$$+\,\eta\big(b_1\cdot\eta(b_2)\cdot b_3\cdot\eta(b_4)\cdot b_5\big)$$



Sixth Moment of Operator-Valued Semicircle

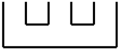
$$E[ssssss] = \eta(1) \cdot \eta(1) \cdot \eta(1)$$

$$+ \eta(1) \cdot \eta\big(\eta(1)\big)$$



$$+ \eta\Big(\eta\big(\eta(1)\big)\Big)$$

$$+ \eta\big(\eta(1)\big) \cdot \eta(1)$$



$$+ \eta\big(\eta(1) \cdot \eta(1)\big)$$

Recursion for Moments of Operator-Valued Semicircle

As before, we have the recurrence relation

$$E[s^{2k}] = \sum_{l=1}^{k} \eta\big(E[s^{2l-2}]\big) \cdot E[s^{2k-2l}].$$

Notation

Put

$$M(z) := \sum_{n=0}^{\infty} E[s^n] z^n = 1 + \sum_{k=1}^{\infty} E[s^{2k}] z^{2k},$$

thus we have again

$$M(z) = 1 + z^2 \eta\big(M(z)\big) \cdot M(z)$$

Recursion for Moments of Operator-Valued Semicircle

$$M(z) = 1 + z^2 \eta\big(M(z)\big) \cdot M(z)$$

Notation (operator-valued Cauchy transform)

Instead of $M(z)$ consider

$$G(z) := E[\frac{1}{z - s}].$$

Note

$$G(z) = E[\frac{1}{z} \cdot \frac{1}{1 - sz^{-1}}] = \frac{1}{z} M(z^{-1}),$$

thus

$$zG(z) = 1 + \eta\big(G(z)\big) \cdot G(z)$$

Thus, the operator-valued Cauchy-transform of s, $G : \mathbb{C}^+ \to \mathcal{B}$, satisfies

$$zG(z) = 1 + \eta\big(G(z)\big) \cdot G(z) \qquad \text{or} \qquad G(z) = \frac{1}{z - \eta\big(G(z)\big)}.$$

This is equivalent to

$$\mathfrak{F}_z(G) = G \qquad \text{where} \qquad \mathfrak{F}_z(G) = \frac{1}{z - \eta(G)}$$

Theorem (Helton, Rashidi Far, Speicher 2007)

For $\Im z > 0$ there exists exactly one solution $G \in \mathbb{H}^-(\mathcal{B})$ to $\mathfrak{F}_z(G) = G$; this G is the limit of iterates $G_n = \mathfrak{F}_z^n(G_0)$ for any $G_0 \in \mathbb{H}^-(\mathcal{B})$. Here

$$H^-(\mathcal{B}) := \{b \in \mathcal{B} \mid \frac{b - b^*}{2i} < 0\}$$

Back to Random Matrices

Basic Observation

Special classes of random matrices are asymptotically described by operator-valued semicircular elements, e.g.

- band matrices (Shlyakhtenko 1996)
- block matrices (Rashidi Far, Oraby, Bryc, Speicher 2006)

Back to Random Matrices

Example

$$X_N = \begin{pmatrix} A_N & B_N & C_N \\ B_N & A_N & B_N \\ C_N & B_N & A_N \end{pmatrix},$$

where A_N, B_N, C_N are independent Gaussian $N \times N$ random matrices.
For $N \to \infty$, X_N converges to

$$s = \begin{pmatrix} s_1 & s_2 & s_3 \\ s_2 & s_1 & s_2 \\ s_3 & s_2 & s_1 \end{pmatrix},$$

where $s_1, s_2, s_3 \in (\mathcal{C}, \varphi)$ is free semicircular family.
This means: the asymptotic eigenvalue distribution of X_N is given by the distribution of s with respect to $\mathrm{tr}_3 \otimes \varphi$.
The latter does not show any nice recursive structure! **But ...**

Example

But $$s = \begin{pmatrix} s_1 & s_2 & s_3 \\ s_2 & s_1 & s_2 \\ s_3 & s_2 & s_1 \end{pmatrix} \qquad (s_1, s_2, s_3 \in (\mathcal{C}, \varphi))$$

is an operator-valued semicircular element over $M_3(\mathbb{C})$ with respect to

- $\mathcal{A} = M_3(\mathcal{C}), \qquad \mathcal{B} = M_3(\mathbb{C})$
- $E = \text{id} \otimes \varphi : M_3(\mathcal{C}) \to M_3(\mathbb{C}), \quad \big(a_{ij}\big)_{i,j=1}^{3} \mapsto \big(\varphi(a_{ij})\big)_{i,j=1}^{3}$
- $\eta : M_3(\mathbb{C}) \to M_3(\mathbb{C})$ given by $\eta(D) = E[sDs]$

Hence asymptotic eigenvalue distribution μ of X_N, which is given by distribution of s with respect to $\text{tr}_3 \otimes \varphi$, can now be factorized as:

$$H(z) = \int \frac{1}{z-t} d\mu(t) = \text{tr}_3 \otimes \varphi\Big(\frac{1}{z-s}\Big) = \text{tr}_3 \Big\{ E\Big[\frac{1}{z-s}\Big] \Big\},$$

and $G(z) = E[\frac{1}{z-s}]$ is solution of $zG(z) = 1 + \eta\big(G(z)\big) \cdot G(z)$

Example

$$s = \begin{pmatrix} s_1 & s_2 & s_3 \\ s_2 & s_1 & s_2 \\ s_3 & s_2 & s_1 \end{pmatrix}: \quad G(z) = \begin{pmatrix} f(z) & 0 & h(z) \\ 0 & g(z) & 0 \\ h(z) & 0 & f(z) \end{pmatrix}, \quad \eta(G) = E[sGs]$$

$$\eta\big(G(z)\big) = \begin{pmatrix} 2\,f(z) + g(z) & 0 & g(z) + 2\,h(z) \\ 0 & 2\,f(z) + g(z) + 2\,h(z) & 0 \\ g(z) + 2\,h(z) & 0 & 2\,f(z) + g(z) \end{pmatrix},$$

$$zG(z) = 1 + \eta\big(G(z)\big) \cdot G(z)$$

$$H(z) = \mathrm{tr}_3(G(z)) = \frac{1}{3}(2f(z) + g(z))$$

System of Quadratic Equations for Operator-Valued Semicircle

Example

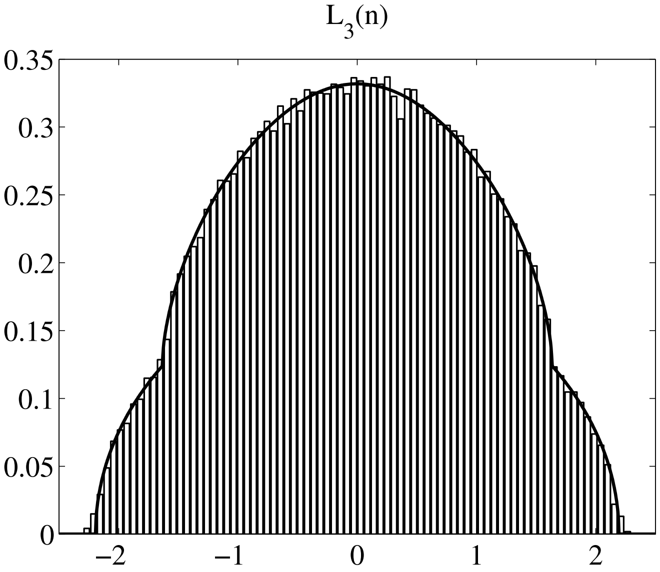
So

$$zG(z) = 1 + \eta\big(G(z)\big) \cdot G(z)$$

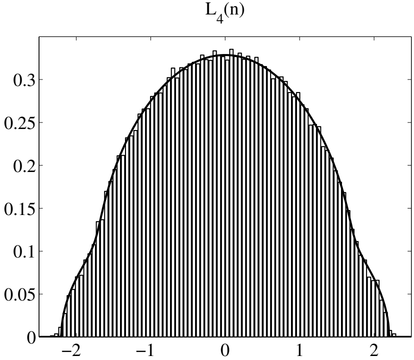
means explicitly

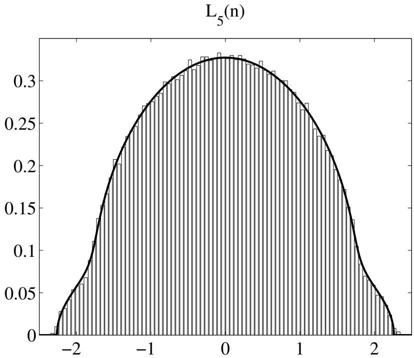
$$\begin{aligned} zf(z) &= 1 + g(z)(f(z) + h(z)) + 2(f(z)^2 + h(z)^2) \\ zg(z) &= 1 + g(z)\big(g(z) + 2(f(z) + h(z))\big) \\ zh(z) &= 4f(z)h(z) + g(z)(f(z) + h(z)) \end{aligned}$$

Comparison of the Solution with Simulations



Some More Examples





$$\begin{pmatrix} A & B & C & D \\ B & A & B & C \\ C & B & A & B \\ D & C & B & A \end{pmatrix}$$

$$\begin{pmatrix} A & B & C & D & E \\ B & A & B & C & D \\ C & B & A & B & C \\ D & C & B & A & B \\ E & D & C & B & A \end{pmatrix},$$

Section 5

Operator-Valued Extension of Free Probability

Problem

What can we say about the relation between two matrices, when we know that the entries of the matrices are free?

$$X = (x_{ij})_{i,j=1}^{N} \qquad Y = (y_{kl})_{k,l=1}^{N}$$

with
$\{x_{ij}\}$ and $\{y_{kl}\}$ free w.r.t. φ

Solution

- X and Y are not free w.r.t. $\mathrm{tr} \otimes \varphi$ in general
- However: relation between X and Y is more complicated, but still treatable in terms of

operator-valued freeness

Notation

Let $(\mathcal{C}, \varphi)$ be non-commutative probability space.
Consider $N \times N$ matrices over $\mathcal{C}$:

$$M_N(\mathcal{C}) := \{(a_{ij})_{i,j=1}^N \mid a_{ij} \in \mathcal{C}\} = M_N(\mathbb{C}) \otimes \mathcal{C}$$

$M_N(\mathcal{C})$ is a non-commutative probability space with respect to

$$\text{tr} \otimes \varphi : M_N(\mathcal{C}) \to \mathbb{C}$$

but there is also an intermediate level

Different Levels

Instead of

$$M_N(\mathcal{C})$$

$$\downarrow \text{tr} \otimes \varphi$$

$$\mathbb{C}$$

consider

$$M_N(\mathcal{C}) = M_N(\mathbb{C}) \otimes \mathcal{C} =: \mathcal{A}$$

$$\downarrow \text{id} \otimes \varphi =: E$$

$$M_N(\mathbb{C}) =: \mathcal{B}$$

$$\downarrow \text{tr}$$

$$\mathbb{C}$$

Definition

Let $\mathcal{B} \subset \mathcal{A}$. A linear map $E : \mathcal{A} \to \mathcal{B}$ is a **conditional expectation** if

$$E[b] = b \qquad \forall b \in \mathcal{B}$$

and

$$E[b_1 a b_2] = b_1 E[a] b_2 \qquad \forall a \in \mathcal{A}, \quad \forall b_1, b_2 \in \mathcal{B}$$

An **operator-valued probability space** consists of $\mathcal{B} \subset \mathcal{A}$ and a conditional expectation $E : \mathcal{A} \to \mathcal{B}$

Example (Classical conditional expectation)

Let $\mathfrak{M}$ be a σ-algebra and $\mathfrak{N} \subset \mathfrak{M}$ be a sub-σ-algebra. Then

- $\mathcal{A} = L^\infty(\Omega, \mathfrak{M}, P)$
- $\mathcal{B} = L^\infty(\Omega, \mathfrak{N}, P)$
- $E[\cdot|\mathfrak{N}]$ is the classical conditional expectation from the bigger onto the smaller σ-algebra.

Example: $M_2(\mathbb{C})$-valued probability space

Example

Let $(\mathcal{A}, \varphi)$ be a non-commutative probability space. Put

$$M_2(\mathcal{A}) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a, b, c, d \in \mathcal{A} \right\}$$

and consider $\psi := \mathrm{tr} \otimes \varphi$ and $E := \mathrm{id} \otimes \varphi$, i.e.:

$$\psi \left[\begin{pmatrix} a & b \\ c & d \end{pmatrix} \right] = \frac{1}{2}(\varphi(a) + \varphi(d)), \qquad E \left[\begin{pmatrix} a & b \\ c & d \end{pmatrix} \right] = \begin{pmatrix} \varphi(a) & \varphi(b) \\ \varphi(c) & \varphi(d) \end{pmatrix}$$

- $(M_2(\mathcal{A}), \psi)$ is a non-commutative probability space, and
- $(M_2(\mathcal{A}), E)$ is an $M_2(\mathbb{C})$-valued probability space

Operator-Valued Distribution

Definition (operator-valued distribution)

Consider an operator-valued probability space $(\mathcal{A}, E : \mathcal{A} \to \mathcal{B})$. The operator-valued distribution of $a \in \mathcal{A}$ is given by all operator-valued moments

$$E[ab_1ab_2 \cdots b_{n-1}a] \in \mathcal{B} \qquad (n \in \mathbb{N}, b_1, \dots, b_{n-1} \in \mathcal{B})$$

Note: polynomials in x with coefficients from $\mathcal{B}$ are of the form

- x^2
- b_0x^2
- $b_1xb_2xb_3$
- $b_1xb_2xb_3 + b_4xb_5xb_6 + \cdots$
- etc.

b's and x do not commute in general!

Definition of Operator-Valued Freeness

Definition (Voiculescu 1985)

Let $E : \mathcal{A} \to \mathcal{B}$ be an operator-valued probability space.
Subalgebras $\mathcal{A}_i$ $(i \in I)$, which contain $\mathcal{B}$, are **free over $\mathcal{B}$**, if
$E[a_1 \cdots a_n] = 0$ whenever

- $a_i \in \mathcal{A}_{j(i)}, \qquad j(i) \in I \quad \forall i$
- $j(1) \neq j(2) \neq \cdots \neq j(n)$
- $E[a_i] = 0 \quad \forall i$

Variables $x_1, \dots, x_n \in \mathcal{A}$ are free over $\mathcal{B}$, if the generated $\mathcal{B}$-subalgebras $\mathcal{A}_i :=$ algebra$(\mathcal{B}, x_i)$ are so.

Freeness and Matrices

Basic Observation

Easy, but crucial fact: Freeness is compatible with going over to matrices

Example

If $\{a_1, b_1, c_1, d_1\}$ and $\{a_2, b_2, c_2, d_2\}$ are free in $(\mathcal{C}, \varphi)$, then

$$\begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix}$$

are

- in general, not free in $(M_2(\mathcal{C}), \text{tr} \otimes \varphi)$
- but free with amalgamation over $M_2(\mathbb{C})$ in $(M_2(\mathcal{C}), \text{id} \otimes \varphi)$

Freeness and Matrices

Example

$$X_1 := \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \qquad \text{and} \qquad X_2 := \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix}$$

Then

$$X_1 X_2 = \begin{pmatrix} a_1a_2 + b_1c_2 & a_1b_2 + b_1d_2 \\ c_1a_2 + d_1c_2 & c_1b_2 + d_1d_2 \end{pmatrix}$$

and

$$\begin{aligned} \psi(X_1X_2) &= \big(\varphi(a_1)\varphi(a_2) + \varphi(b_1)\varphi(c_2) + \varphi(c_1)\varphi(b_2) + \varphi(d_1)\varphi(d_2)\big)/2 \\ &\neq (\varphi(a_1) + \varphi(d_1))(\varphi(a_2) + \varphi(d_2))/4 \\ &= \psi(X_1) \cdot \psi(X_2) \end{aligned}$$

but

$$E(X_1X_2) = E(X_1) \cdot E(X_2)$$

Freeness and Matrices

Example

$$X_1 := \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \qquad \text{and} \qquad X_2 := \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix}$$

Then

$$X_1 X_2 = \begin{pmatrix} a_1 a_2 + b_1 c_2 & a_1 b_2 + b_1 d_2 \\ c_1 a_2 + d_1 c_2 & c_1 b_2 + d_1 d_2 \end{pmatrix}$$

and

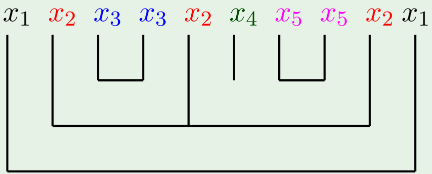
$$\begin{aligned} E(X_1 X_2) &= \begin{pmatrix} \varphi(a_1)\varphi(a_2) + \varphi(b_1)\varphi(c_2) & \varphi(a_1)\varphi(b_2) + \varphi(b_1)\varphi(d_2) \\ \varphi(c_1)\varphi(a_2) + \varphi(d_1)\varphi(c_2) & \varphi(c_1)\varphi(b_2) + \varphi(d_1)\varphi(d_2) \end{pmatrix} \\ &= \begin{pmatrix} \varphi(a_1) & \varphi(b_1) \\ \varphi(c_1) & \varphi(d_1) \end{pmatrix} \begin{pmatrix} \varphi(a_2) & \varphi(b_2) \\ \varphi(c_2) & \varphi(d_2) \end{pmatrix} \\ &= E(X_1) \cdot E(X_2) \end{aligned}$$

Combinatorial Description of Operator-Valued Freeness

Operator-valued freeness works mostly like ordinary freeness, one only has to take care of the order of the variables; in all expressions they have to appear in their original order!

Example

Still one has factorizations of all non-crossing moments in free variables.



$$E[x_1 x_2 x_3 x_3 x_2 x_4 x_5 x_5 x_2 x_1]$$
$$= E\Big[x_1 \cdot E\big[x_2 \cdot E[x_3 x_3] \cdot x_2 \cdot E[x_4] \cdot E[x_5 x_5] \cdot x_2\big] \cdot x_1\Big]$$

Combinatorial Description of Operator-Valued Freeness

For “crossing" moments one has analogous formulas as in scalar-valued case, modulo respecting the order of the variables ...

Example

The formula

$$\varphi(x_1x_2x_1x_2) = \varphi(x_1x_1)\varphi(x_2)\varphi(x_2) + \varphi(x_1)\varphi(x_1)\varphi(x_2x_2) - \varphi(x_1)\varphi(x_2)\varphi(x_1)\varphi(x_2)$$

has now to be written as

$$E[x_1x_2x_1x_2] = E\big[x_1E[x_2]x_1\big] \cdot E[x_2] + E[x_1] \cdot E\big[x_2E[x_1]x_2\big] - E[x_1]E[x_2]E[x_1]E[x_2]$$

Free Cumulants

Definition

Consider $E : \mathcal{A} \to \mathcal{B}$.
Define **free cumulants**

$$k_n^{\mathcal{B}} : \mathcal{A}^n \to \mathcal{B}$$

by

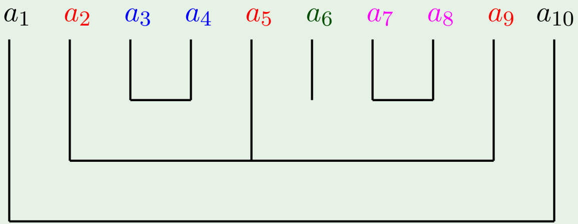
$$E[a_1 \cdots a_n] = \sum_{\pi \in NC(n)} k_\pi^{\mathcal{B}}[a_1, \dots, a_n]$$

- arguments of $k_\pi^{\mathcal{B}}$ are distributed according to blocks of π
- but now: cumulants are nested inside each other according to nesting of blocks of π

Free Cumulants

Example

$$\pi = \big\{\{1,10\},\{2,5,9\},\{3,4\},\{6\},\{7,8\}\big\} \in NC(10),$$



$k_\pi^{\mathcal{B}}[a_1,\dots,a_{10}]$

$$= k_2^{\mathcal{B}}\Big(a_1 \cdot k_3^{\mathcal{B}}\big(a_2 \cdot k_2^{\mathcal{B}}(a_3,a_4), a_5 \cdot k_1^{\mathcal{B}}(a_6) \cdot k_2^{\mathcal{B}}(a_7,a_8), a_9\big), a_{10}\Big)$$

Analytic Description of Operator-Valued Free Convolution

Definition

For a random variable $x \in \mathcal{A}$ in an operator-valued probability space $E : \mathcal{A} \to \mathcal{B}$ we define the **operator-valued Cauchy transform**:

$$G(b) := E[(b - x)^{-1}] \qquad (b \in \mathcal{B}).$$

For $x = x^*$, this is well-defined and a nice analytic map on the

operator-valued upper halfplane $\qquad \mathbb{H}^+(\mathcal{B}) := \{b \in \mathcal{B} \mid \frac{b - b^*}{2i} > 0\}$

Definition

We define the **operator-valued R-transform** by

$$bG(b) = 1 + R(G(b)) \cdot G(b) \qquad \text{or} \qquad G(b) = \frac{1}{b - R(G(b))}$$

On a Formal Power Series Level: Same Results as in Scalar-Valued Case

Note that for an operator-valued semicircular element with covariance η we have $R(b) = \eta(b)$ and thus

$$bG(b) = 1 + R(G(b)) \cdot G(b), \qquad \text{restricted to } b = z,$$

is nothing but our formula from before

$$zG(z) = 1 + \eta\big(G(z)\big) \cdot G(z)$$

If x and y are free over $\mathcal{B}$, then

- mixed $\mathcal{B}$-valued cumulants in x and y vanish
- $R_{x+y}(b) = R_x(b) + R_y(b)$
- we have the subordination $G_{x+y}(z) = G_x(\omega(z))$

Subordination in the Operator-Valued Case

- again, analytic properties of R transform are not so nice
- the operator-valued equation $G(b) = \frac{1}{b-R(G(b))}$, has hardly ever explicit solutions and, from the numerical point of view, it becomes quite intractable: instead of one algebraic equation we have now a system of algebraic equations
- subordination version for the operator-valued case was treated by Biane (1998) and, more conceptually, by Voiculescu (2000)
- an analytic description of subordination via fixed point equations, as in the scalar-valued case, was given by Belinschi, Mai, Speicher (2013)

Subordination Formulation

Theorem (Belinschi, Mai, Speicher 2013)

Let x and y be selfadjoint operator-valued random variables free over $\mathcal{B}$. Then there exists a Fréchet analytic map $\omega\colon \mathbb{H}^+(\mathcal{B}) \to \mathbb{H}^+(\mathcal{B})$ so that

$$G_{x+y}(b) = G_x(\omega(b)) \text{ for all } b \in \mathbb{H}^+(\mathcal{B}).$$

Moreover, if $b \in \mathbb{H}^+(\mathcal{B})$, then $\omega(b)$ is the unique fixed point of the map

$$f_b\colon \mathbb{H}^+(\mathcal{B}) \to \mathbb{H}^+(\mathcal{B}), \quad f_b(w) = h_y(h_x(w) + b) + b,$$

and

$$\omega(b) = \lim_{n\to\infty} f_b^{\circ n}(w) \qquad \text{for any } w \in \mathbb{H}^+(\mathcal{B}).$$

where

$$\mathbb{H}^+(\mathcal{B}) := \{b \in \mathcal{B} \mid \frac{b - b^*}{2i} > 0\}, \qquad h(b) := \frac{1}{G(b)} - b$$

Section 6

Deterministic Equivalents

Problem

Quite often, one has a random matrix problem for (large) size N, but the limit $N \to \infty$ is not adequate, because there is no canonical limit for some of the involved matrices

Solution (Girko; Couillet, Hoydis, Debbah; Hachem, Loubaton, Najim)

Deterministic Equivalent: Replace the random Stieltjes transform g_N of the problem for N by a deterministic transform $\tilde{g}_N$ such that

- $\tilde{g}_N$ is calculable, usually as the fixed point solution of some system of equations
- the difference between g_N and $\tilde{g}_N$ goes, for $N \to \infty$, to 0 (even though g_N itself might not converge)

Deterministic Equivalent

- Replace the original unsolvable problem by another problem which is
  - solvable
  - close to the original problem (at least for large N)
- The replacement is done on the level of Stieltjes transforms and there is no clear rule how to do this
- Essentially one tries to close the system of equations for the Stieltjes transforms by keeping as much data as possible of the original situation
- Replacement and solving is done in one step

Free Deterministic Equivalent (Speicher, Vargas)

- We will replace the original problem by another one on the level of operators in a quite precise way, essentially by prescribing
  - replace Gaussian random matrices by semicircular variables
  - replace matrices which are asymptotically free by free variables
- The free deterministic equivalent is then a well-defined function in free variables
- That the free deterministic equivalent is close to the original model (for large N) is essentially the same calculation as showing asymptotic freeness
- One can then try to solve for the distribution of this replacement in a second step

Free Deterministic Equivalent (Speicher, Vargas)

Example

Consider $A_N = T_N + X_N$ where

- X_N is a symmetric $N \times N$ Gaussian random matrix
- T_N is a deterministic matrix

We do not have a sequence T_N, with $N \to \infty$, thus we only have the distribution of T_N for some fixed N.
We replace now A_N by $a_N = t_N + s$, where

- s is a semicircular element
- t_N is an operator which has the same distribution as T_N
- t_N and s are free

In this case, the distribution of a_N is given by the free convolution of the distribution of t_N and the distribution of s,

$$\mu_{A_N} \sim \mu_{a_N} = \mu_{t_N+s} = \mu_{t_N} \boxplus \mu_s = \mu_{T_N} \boxplus \mu_s$$

Can We Calculate Free Deterministic Equivalents?

Problem

Usually, our free deterministic equivalents are polynomials in free variables. Can we calculate their distribution out of the knowledge of the distribution of each variable?

Solution

Yes, we can!

For this, use the combination of

- the linearization trick
- and recent advances on the analytic description of operator-valued free convolution

Section 7

The Linearization Trick

The Linearization Philosophy

In order to understand polynomials in non-commuting variables, it suffices to understand matrices of **linear** polynomials in those variables.

History (in operator algebras)

- Voiculescu 1987: motivation
- Haagerup, Thorbjørnsen 2005: largest eigenvalue
- Anderson 2012: the selfadjoint version ("Schur complement")

History (in other fields)

The same idea has been used in other fields under different names (like "descriptor system" in control theory), for example:

- Schützenberger 1961: automata theory
- Helton, McCullough, Vinnikov 2006: symmetric descriptor realization

Definition

Consider a polynomial p in non-commuting variables x and y.
A **linearization** of p is an $N \times N$ matrix (with $N \in \mathbb{N}$) of the form

$$\hat{p} = \begin{pmatrix} 0 & u \\ v & Q \end{pmatrix},$$

- u, v, Q are matrices of the following sizes: u is $1 \times (N-1)$; v is $(N-1) \times 1$; and Q is $(N-1) \times (N-1)$
- u, v, Q are polynomials in x and y, each of degree ≤ 1
- Q is invertible and we have $p = -uQ^{-1}v$

Theorem (Schützenberger; Helton, McCullough, Vinnikov; Anderson)

- For each p there exists a linearization $\hat{p}$ (with an explicit algorithm for finding those)
- If p is selfadjoint, then this $\hat{p}$ is also selfadjoint

Theorem (Schützenberger; Helton, McCullough, Vinnikov; Anderson)

- For each p there exists a linearization $\hat{p}$ (with an explicit algorithm for finding those)
- If p is selfadjoint, then this $\hat{p}$ is also selfadjoint

Example

A selfadjoint linearization of

$$p = xy + yx + x^2 \qquad \text{is} \qquad \hat{p} = \begin{pmatrix} 0 & x & \frac{x}{2}+y \\ x & 0 & -1 \\ \frac{x}{2}+y & -1 & 0 \end{pmatrix}$$

because we have

$$\begin{pmatrix} x & \frac{x}{2}+y \end{pmatrix} \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}^{-1} \begin{pmatrix} x \\ \frac{x}{2}+y \end{pmatrix} = -(xy + yx + x^2)$$

What is a Linearization Good for?

We have then

$$\hat{p} = \begin{pmatrix} 0 & u \\ v & Q \end{pmatrix} = \begin{pmatrix} 1 & uQ^{-1} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} p & 0 \\ 0 & Q \end{pmatrix} \begin{pmatrix} 1 & 0 \\ Q^{-1}v & 1 \end{pmatrix}$$

and thus (under the condition that Q is invertible):

$$p \text{ invertible} \qquad \Longleftrightarrow \qquad \hat{p} \text{ invertible}$$

Note: $\begin{pmatrix} 1 & 0 \\ a & 1 \end{pmatrix}$ is always invertible with

$$\begin{pmatrix} 1 & 0 \\ a & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 \\ -a & 1 \end{pmatrix}$$

What is a Linearization Good for?

More general, for $z \in \mathbb{C}$ put $b = \begin{pmatrix} z & 0 \\ 0 & 0 \end{pmatrix}$ and then

$$b - \hat{p} = \begin{pmatrix} z & -u \\ -v & -Q \end{pmatrix} = \begin{pmatrix} 1 & uQ^{-1} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} z - p & 0 \\ 0 & -Q \end{pmatrix} \begin{pmatrix} 1 & 0 \\ Q^{-1}v & 1 \end{pmatrix}$$

$$z - p \text{ invertible} \quad \Longleftrightarrow \quad b - \hat{p} \text{ invertible}$$

and actually

$$(b - \hat{p})^{-1} = \left[\begin{pmatrix} 1 & uQ^{-1} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} z - p & 0 \\ 0 & -Q \end{pmatrix} \begin{pmatrix} 1 & 0 \\ Q^{-1}v & 1 \end{pmatrix} \right]^{-1}$$

$$= \begin{pmatrix} 1 & 0 \\ -Q^{-1}v & 1 \end{pmatrix} \begin{pmatrix} (z - p)^{-1} & 0 \\ 0 & -Q^{-1} \end{pmatrix} \begin{pmatrix} 1 & -uQ^{-1} \\ 0 & 1 \end{pmatrix}$$

$$(b-\hat{p})^{-1} = \begin{pmatrix} 1 & 0 \\ -Q^{-1}v & 1 \end{pmatrix} \begin{pmatrix} (z-p)^{-1} & 0 \\ 0 & -Q^{-1} \end{pmatrix} \begin{pmatrix} 1 & -uQ^{-1} \\ 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} (z-p)^{-1} & -(z-p)^{-1}uQ^{-1} \\ -Q^{-1}v(z-p)^{-1} & Q^{-1}v(z-p)^{-1}uQ^{-1} - Q^{-1} \end{pmatrix}$$

$$= \begin{pmatrix} (z-p)^{-1} & * \\ * & * \end{pmatrix}$$

and we can get the Cauchy transform $G_p(z) = \varphi((z-p)^{-1})$ of p as the (1,1)-entry of the matrix-valued Cauchy-transform of $\hat{p}$

$$G_{\hat{p}}(b) = \text{id} \otimes \varphi((b-\hat{p})^{-1}) = \begin{pmatrix} \varphi((z-p)^{-1}) & \cdots \\ \cdots & \cdots \end{pmatrix}$$

Why is $\hat{p}$ better than p?

The selfadjoint linearization $\hat{p}$ is now the sum of two selfadjoint operator-valued variables

$$\hat{p} = \hat{x} + \hat{y} = \begin{pmatrix} 0 & x & \frac{x}{2} \\ x & 0 & 0 \\ \frac{x}{2} & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & y \\ 0 & 0 & -1 \\ y & -1 & 0 \end{pmatrix}$$

where

- we know the operator-valued distribution of $\hat{x}$ and the operator-valued distribution of $\hat{y}$
- and $\hat{x}$ and $\hat{y}$ are operator-valued freely independent!

This is now a problem about operator-valued free convolution. This we can do.

The selfadjoint linearization $\hat{p}$ is now the sum of two selfadjoint operator-valued variables

$$\hat{p} = \hat{x} + \hat{y} = \begin{pmatrix} 0 & x & \frac{x}{2} \\ x & 0 & 0 \\ \frac{x}{2} & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & y \\ 0 & 0 & -1 \\ y & -1 & 0 \end{pmatrix}$$

where

- we know the operator-valued distribution of $\hat{x}$ and the operator-valued distribution of $\hat{y}$
- and $\hat{x}$ and $\hat{y}$ are operator-valued freely independent!

So we can use operator-valued free convolution to calculate the operator-valued Cauchy transform of $\hat{x} + \hat{y}$.

$$G_{\hat{p}}(b) = G_{\hat{x}}(\omega(b))$$

and from this get the Caucy transform of $p(x, y)$.

Theorem (Belinschi, Mai, Speicher 2013)

1) The following algorithm allows the calculation of the distribution of any selfadjoint polynomial $p(x, y)$ in two free variables x and y, given the distribution of x and the distribution of y:

- *Linearize $p(x, y)$ to $\hat{p} = \hat{x} + \hat{y}$.*
- *Calculate $G_{\hat{x}}(b)$ out of $G_x(z)$ and $G_{\hat{y}}(b)$ out of $G_y(z)$*
- *Get $w_1(b)$ as the fixed point of the iteration*

$$w \mapsto G_{\hat{y}}(b + G_{\hat{x}}(w)^{-1} - w)^{-1} - (G_{\hat{x}}(w)^{-1} - w)$$

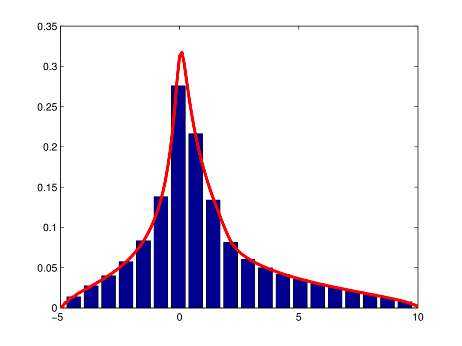
- *Calculate $G_{\hat{p}}(b) = G_{\hat{x}}(\omega_1(b))$ and recover $G_p(z)$ as one entry of $G_{\hat{p}}(b)$ for $b = \begin{pmatrix} z & 0 \\ 0 & 0 \end{pmatrix}$*

2) Iteration of step 3 of the above algorithm allows the calculation of the distribution of any selfadjoint polynomial $p(x_1, \ldots, x_k)$ in k non-commuting variables, given the distribution of each x_i.

Example

$P(X,Y) = XY + YX + X^2$
for independent X, Y; X is Gaussian and Y is Wishart

$$\hat{p} = \begin{pmatrix} 0 & x & y + \frac{x}{2} \\ x & 0 & -1 \\ y + \frac{x}{2} & -1 & 0 \end{pmatrix}$$

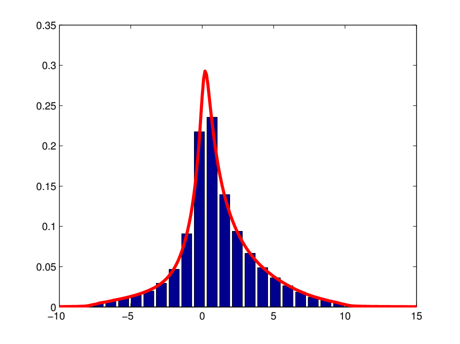


$p(x,y) = xy + yx + x^2$
for free x, y; x is semicircular and y is Marchenko-Pastur

Example

$P(X_1, X_2, X_3) = X_1X_2X_1 + X_2X_3X_2 + X_3X_1X_3$
for independent X_1, X_2, X_3; X_1, X_2 Wigner, X_3 Wishart

$$\hat{p} = \begin{pmatrix} 0 & 0 & x_1 & 0 & x_2 & 0 & x_3 \\ 0 & x_2 & -1 & 0 & 0 & 0 & 0 \\ x_1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_3 & -1 & 0 & 0 \\ x_2 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & x_1 & -1 \\ x_3 & 0 & 0 & 0 & 0 & -1 & 0 \end{pmatrix}$$



$p(x_1, x_2, x_3) = x_1x_2x_1 + x_2x_3x_2 + x_3x_1x_3$
for free x_1, x_2, x_3; x_1, x_2 is semicircular and x_3 is Marchenko-Pastur

A Bit on the Linearization Algorithm

Problem

We want to find selfadjoint linearization a non-commutative polynomial p. For this consider the following steps.

1. Calculate a linearization for each monomial of p.
2. Given linearizations of monomials $q_1, \ldots, q_n$, what is a linearization of $q_1 + \ldots + q_n$?
3. Consider a polynomial p of the form $q + q^*$ and let $\hat{q}$ be a linarization of q. Calculate a linearization of p in terms of $\hat{q}$.

Solution

1. A linearization of $q = x_i x_j x_k$ is
$$\hat{q} = \begin{pmatrix} 0 & 0 & x_i \\ 0 & x_j & -1 \\ x_k & -1 & 0 \end{pmatrix}.$$

2. We consider two linearizations $\hat{q}_1 = \begin{pmatrix} 0 & u_1 \\ v_1 & Q_1 \end{pmatrix}$ and $\hat{q}_2 = \begin{pmatrix} 0 & u_2 \\ v_2 & Q_2 \end{pmatrix}$.
A linearization $\widehat{q_1 + q_2}$ of $q_1 + q_2$ is given by
$$\begin{pmatrix} 0 & u_1 & u_2 \\ v_1 & Q_1 & 0 \\ v_2 & 0 & Q_2 \end{pmatrix}.$$

3. If $\hat{q} = \begin{pmatrix} 0 & u \\ v & Q \end{pmatrix}$ then we can choose $\widehat{q + q^*} = \begin{pmatrix} 0 & u & v^* \\ u^* & 0 & Q \\ v & Q^* & 0 \end{pmatrix}$.

Solution

1. Linearizations of $x_1x_2x_1$, $x_2x_3x_2$, $x_3x_1x_3$ are

$$\begin{pmatrix} 0 & 0 & x_1 \\ 0 & x_2 & -1 \\ x_1 & -1 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 0 & x_2 \\ 0 & x_3 & -1 \\ x_2 & -1 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 0 & x_3 \\ 0 & x_1 & -1 \\ x_3 & -1 & 0 \end{pmatrix}$$

2. thus a linearization of $p(x_1, x_2, x_3) = x_1x_2x_1 + x_2x_3x_2 + x_3x_1x_3$ is

$$\begin{pmatrix} 0 & 0 & x_1 & 0 & x_2 & 0 & x_3 \\ 0 & x_2 & -1 & 0 & 0 & 0 & 0 \\ x_1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_3 & -1 & 0 & 0 \\ x_2 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & x_1 & -1 \\ x_3 & 0 & 0 & 0 & 0 & -1 & 0 \end{pmatrix}.$$