Permutation Invariant Gaussian Matrix Models

George Barnes

Queen Mary University of London

March 2021

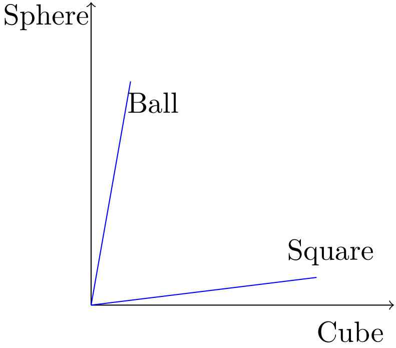
- Motivation: Computational lingustics
- Permutation invariant Gaussian matrix model
- Experimental test
- Extending the model

Motivation: Linguistics

"You shall know a word by the company it keeps" (Firth 1957)

Distributional hypothesis: The meaning of a word can be represented by a vector recording the frequency of its cooccurrence with other words. (Harris 1954)

The basis of this vector space is a set of commonly occurring "context words".



Distributional semantics works well for nouns but has problems with anything more complex

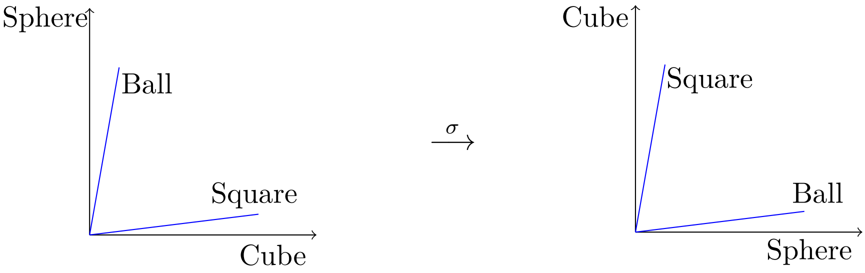
To overcome this difficulty recent work has focussed on compositional models of meaning (Coecke, Sadrzadeh, Clark '10)

These models introduce higher index objects in order to represent grammatical structure

$$(red)_{ij}(box)_j = (red\ box)_i,$$
$$(like)_{ijk}(cats)_j(fish)_k = (cats\ like\ fish)_i$$

Expect the meaning of a word to be independent of the ordering of the context words

A permutation of the basis vectors does not change the overlap of a meaning vector with any given context word



There is evidence that the components of a word's meaning representation follow a Gaussian distribution

(Kartsaklis, Ramgoolam, Sadrzadeh '17)

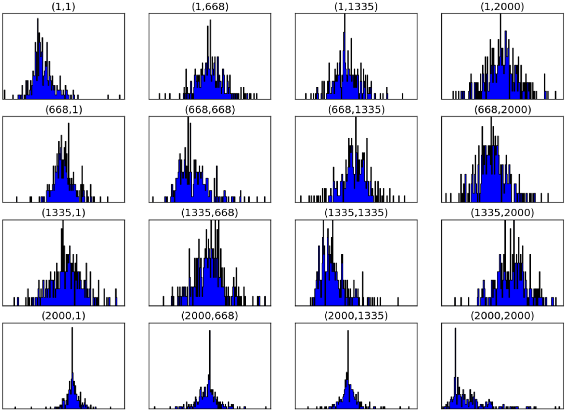


Figure 1: Histograms for elements of adjective matrices

Permutation invariant Gaussian matrix model

Permutation invariant Gaussian matrix model

What do these models look like?

Zero dimensional QFT with matrix valued fields

$$\mathcal{Z} = \int dM e^{-\mathcal{S}(M)}$$

Gaussian $\rightarrow$ No interaction terms

$$\mathcal{S}(M) \sim M + M^2$$

Permutation invariant $\rightarrow$ The action of the theory is unchanged by permuting the matrix elements

$$\mathcal{S}(M_{ij}) = \mathcal{S}(M_{\sigma(i)\sigma(j)}), \qquad \sigma \in S_D$$

The observables of interest are the permutation invariant polynomials

$$f(M_{ij}) = f(M_{\sigma(i)\sigma(j)}), \qquad \forall \sigma \in S_D$$

We can calculate expectation values via

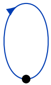
$$\langle f(M_{ij}) \rangle = \frac{1}{\mathcal{Z}} \int dM e^{-\mathcal{S}(M)} f(M_{ij})$$

Permutation invariant Gaussian matrix model

Observable graph correspondence

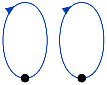
There is a 1-1 correspondence between these observables and directed graphs

Each unique index is associated with a vertex and each matrix is associated with a directed edge from the first of its indices to the second

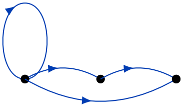


$$\Leftrightarrow \sum_{i} M_{ii},$$

$$\Leftrightarrow \sum_{i,j} M_{ij}$$



$$\Leftrightarrow \sum_{i,j} M_{ii} M_{jj},$$



$$\Leftrightarrow \sum_{i,j,k} M_{ii} M_{ij} M_{jk} M_{ik}$$

Permutation invariant Gaussian matrix model

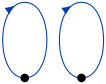
Observable graph correspondence

There are 11 invariants at quadratic order:

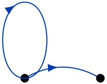
$$\sum_{i} M_{ii}M_{ii}$$

$$\sum_{i,j} M_{ij}M_{ij}$$

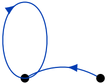
$$\sum_{i,j} M_{ij}M_{ji}$$



$$\sum_{i,j} M_{ii}M_{jj}$$



$$\sum_{i,j} M_{ii}M_{ij}$$

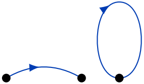


$$\sum_{i,j} M_{ii}M_{ji}$$

$$\sum_{i,j,k} M_{ij}M_{jk}$$

$$\sum_{i,j,k} M_{ij}M_{ik}$$

$$\sum_{i,j,k} M_{ij}M_{kj}$$



$$\sum_{i,j,k} M_{ij}M_{kk}$$

$$\sum_{i,j,k,l} M_{ij}M_{kl}$$

We can define a five-parameter model with the partition function

$$\mathcal{Z}(\Lambda, a, b, J^0, J^S) = \int dM e^{J^0 \sum_{i=1}^{D} M_{ii} + J^S \sum_{i<j}(M_{ij}+M_{ji}) - \frac{\Lambda}{2}\sum_i M_{ii}^2} e^{-\frac{1}{4}(a+b)\sum_{i<j}(M_{ij}^2+M_{ji}^2) - \frac{1}{2}(a-b)\sum_{i<j} M_{ij}M_{ji}}$$

This factorises into D integrals for the diagonal matrix elements and $D(D-1)/2$ integrals for the off-diagonal elements

This model with two linear terms and three quadratic is solvable but not the most general (Kartsaklis, Ramgoolam, Sadrzadeh '17)

The most general Gaussian action is comprised of a weighted sum of all possible linear and quadratic terms

$$\sum_{i=1}^{D} M_{ii}, \quad \sum_{i,j=1}^{D} M_{ij}, \quad \sum_{i=1}^{D} M_{ii}^2, \quad \sum_{i,j=1}^{D} M_{ij} M_{ji}, \ldots, \sum_{i,j,k,l=1}^{D} M_{ij} M_{kl}$$

This action mixes the D^2 elements M_{ij} in some complicated way such that we are not able to solve the partition function in this form

$$\mathcal{Z} = \int dM \exp\Big(-\frac{1}{2} x_\alpha \Lambda_{\alpha\beta} x_\beta + \rho_\alpha x_\alpha \Big) = \sqrt{\frac{(2\pi)^N}{\det\Lambda}} \exp\Big(\frac{1}{2} \rho_\alpha (\Lambda^{-1})_{\alpha\beta} \rho_\beta \Big)$$

It is possible to solve the most general Gaussian one-matrix action (Ramgoolam '18)

Look for a change of variables that factorises the partition function (at least into block diagonal form) $\Rightarrow$ representation theory of the symmetric group

V_D is the natural representation of the symmetric group on D symbols. Consider it as the span of D orthonormal basis vectors $\{e_1, e_2, \ldots, e_D\}$ with the action of $\sigma \in S_D$ given by

$$\rho_{V_D}(\sigma)e_i = e_{\sigma^{-1}(i)}$$

and extended by linearity.

The D^2 matrix elements M_{ij} transform as the product of two copies of the natural representation $V_D \otimes V_D$

V_D is a reducible representation of the symmetric group. There is an invariant vector in this space given by

$$E_0 = \frac{1}{\sqrt{D}} \sum_{i=1}^{D} e_i$$

The following $D-1$ linear combinations

$$E_1 = \frac{1}{\sqrt{2}}(e_1 - e_2),$$
$$E_2 = \frac{1}{\sqrt{6}}(e_1 + e_2 - 2e_3),$$
$$\vdots$$
$$E_a = \frac{1}{\sqrt{a(a+1)}}(e_1 + e_2 + \cdots + e_a - ae_{a+1})$$

with $1 \leq a \leq D-1$ form an S_D-invariant subspace of V_D. The E_a form a basis of an irrep of the symmetric group called the Hook representation.

The natural representation of the symmetric group irreducibly decomposes as

$$V_D = V_0 \oplus V_H$$

We would like to find the transformation that reduces $V_D \otimes V_D$ to a direct sum of irreducible representations of $\mathrm{Diag}(S_D)$.

$$\begin{aligned} M_{ij} &\cong V_D \otimes V_D \\ &= (V_0 \oplus V_H) \otimes (V_0 \oplus V_H) \\ &= (V_0 \otimes V_0) \oplus (V_0 \otimes V_H) \oplus (V_H \otimes V_0) \oplus (V_H \otimes V_H) \end{aligned}$$

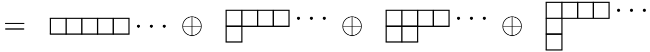
The only term that transforms as a reducible representation of S_D is $V_H \otimes V_H$.

Permutation invariant Gaussian matrix model

Representation theory

The $V_H \otimes V_H$ space can be decomposed into irreps of the diagonal S_D as

$$V_H \otimes V_H = V_0 \quad \oplus V_H \quad \oplus V_2 \quad \oplus V_3$$



Leaving us with

$$\text{Span}\{M_{ij} : 1 \leq i, j \leq D\} = \bigoplus_{\alpha=1}^{2} V_0^{\alpha} \bigoplus_{\alpha=1}^{3} V_H^{\alpha} \oplus V_2 \oplus V_3$$

We define the following variables transforming according to this decomposition

Trivial rep:	$S^{00}, S^{HH\to 0}$	$S^{V_0;\alpha}, \quad \alpha = \{1,2\}$
Hook rep:	$S^{0H}_a, S^{H0}_a, S^{HH\to H}_a$	$S^{V_H;\alpha}_a, \quad \alpha = \{1,2,3\}$
The rep V2:	$S^{HH\to V_2}_a$	$S^{V_2}_a$
The rep V3:	$S^{HH\to V_3}_a$	$S^{V_3}_a$

where superscripts track the representation theoretic origin of each variable.

The orthonormality of the e_i basis of V_D

$$\langle e_i | e_j \rangle = \delta_{ij}, \quad 1 \leq i, j \leq D.$$

implies the same for E_A

$$\langle E_A | E_B \rangle = \delta_{AB}, \quad 0 \leq A, B \leq D - 1.$$

We then define the overlap coefficients

$$C_{0,i} = \langle E_0 | e_i \rangle = \frac{1}{\sqrt{D}}$$

$$C_{a,i} = \langle E_a | e_i \rangle = \frac{1}{\sqrt{a(a+1)}} \Big(\sum_{j=1}^{a} \delta_{ij} - a \delta_{i,a+1} \Big)$$

Using the orthonormality we use

$$\sum_{A=0}^{D-1} C_{A,i} C_{A,j} = C_{0,i} C_{0,j} + \sum_{a=1}^{D-1} C_{a,i} C_{a,j} = \delta_{ij}$$

to define an object of central importance in the calculation of correlators

$$\sum_{a=1}^{D-1} C_{a,i} C_{a,j} = \delta_{ij} - \frac{1}{D} \equiv F(i,j)$$

This is the projector in V_D for the V_H subspace.

Indeed $F(i,j)$ behaves like a projector obeying

$$\sum_{j=1}^{D} F(i,j)F(j,k) = F(i,k)$$

and

$$\sum_{i=1}^{D} F(i,i) = D-1$$

Furthermore

$$\sum_{i=1}^{D} F(i,j) = 0$$

It is possible to write the Clebsch-Gordan coefficients in terms of the $C_{a,i}$ and $C_{0,i}$

$$
\begin{aligned}
M_{ij} = & \frac{1}{D} S^{00} + \frac{1}{\sqrt{D-1}} \sum_{a=1}^{D-1} C_{a,i} C_{a,j} S^{HH \to 0} + \frac{1}{\sqrt{D}} \sum_{a=1}^{D-1} C_{a,j} S_a^{0H} \\
& + \frac{1}{\sqrt{D}} \sum_{a=1}^{D-1} C_{a,i} S_a^{H0} + \sum_{a,b,c=1}^{D-1} C_{a,i} C_{b,j} C_{a,b;\ \ c}^{V_H V_H \to V_H} S_c^{HH \to H} \\
& + \sum_{a,b=1}^{D-1} \sum_{c=1}^{\mathrm{Dim} V_2} C_{a,i} C_{b,j} C_{a,b;\ \ c}^{V_H V_H \to V_2} S_c^{HH \to V_2} \\
& + \sum_{a,b=1}^{D-1} \sum_{c=1}^{\mathrm{Dim} V_3} C_{a,i} C_{b,j} C_{a,b;\ \ c}^{V_H V_H \to V_3} S_c^{HH \to V_3} .
\end{aligned}
$$

Our action should also include terms quadratic in M_{ij}

Every irreducible representation of S_D, call it V_R has the property that

$$\mathrm{Sym}^2(V_R)$$

contains the trivial representation exactly once

The square of M transforms as

$$\begin{aligned} M^2 &\cong \mathrm{Sym}^2(V_D \otimes V_D) \\ &= \mathrm{Sym}^2(2V_0 \oplus 3V_H \oplus V_2 \oplus V_3) \end{aligned}$$

There are two copies of V_0 in this decomposition: V_0^{00} and V_0^{HH}

$\mathrm{Sym}^2(V_0^{00} \oplus V_0^{HH})$ contains three invariants:

$$(S^{00})^2, \quad S^{00} S^{HH\to 0} = S^{HH\to 0} S^{00}, \quad (S^{HH\to 0})^2$$

The general invariant quadratic function of the variables transforming as V_0 is

$$\sum_{\alpha,\beta=1}^{2} (\Lambda_{V_0})_{\alpha\beta} S^{V_0;\alpha} S^{V_0;\beta}$$

with Λ_{V_0} a 2×2 symmetric matrix.

Similarly, for the V_H, V_2 and V_3 variables we have

$$\sum_{\alpha,\beta=1}^{3} (\Lambda_{V_H})_{\alpha\beta} \sum_{a}^{D-1} S_a^{V_H;\alpha} S_a^{V_H;\beta}$$

$$\Lambda_{V_2} \sum_{a}^{\mathrm{Dim} V_2} S_a^{V_2} S_a^{V_2}$$

$$\Lambda_{V_3} \sum_{a}^{\mathrm{Dim} V_2} S_a^{V_3} S_a^{V_3}$$

respectively, with Λ_{V_H} a 3×3 symmetric matrix, and Λ_{V_2}, Λ_{V_3} numbers.

We now have a partition function

$$\mathcal{Z}(\rho_1, \rho_2; \Lambda_{V_0}, \Lambda_{V_H}, \Lambda_{V_2}, \Lambda_{V_3}) = \int dM e^{-\mathcal{S}}$$

Our action can now be written

$$\begin{aligned}\mathcal{S} = &- \sum_{\alpha=1}^{2} \rho_\alpha^{V_0} S^{V_0;\alpha} + \frac{1}{2} \sum_{\alpha,\beta=1}^{3} S^{V_0;\alpha} (\Lambda_{V_0})_{\alpha\beta} S^{V_0;\beta} \\ &+ \frac{1}{2} \sum_{a=1}^{D-1} \sum_{\alpha,\beta=1}^{3} S_a^{V_H;\alpha} (\Lambda_{V_H})_{\alpha\beta} S_a^{V_H;\beta} + \frac{1}{2} \Lambda_{V_2} \sum_{a=1}^{\mathrm{Dim} V_2} S_a^{V_2} S_a^{V_2} \\ &+ \frac{1}{2} \Lambda_{V_3} \sum_{a=1}^{\mathrm{Dim} V_3} S_a^{V_3} S_a^{V_3}\end{aligned}$$

The measure dM for integration over the matrix variables M_{ij} is taken to be the Euclidean measure

$$dM = \prod_i dM_{ii} \prod_{i \neq j} dM_{ij}$$

It can be shown that

$$dM = dS^{V_0;1} dS^{V_0;2} \prod_{a=1}^{D-1} dS_a^{V_H;1} dS_a^{V_H;2} dS_a^{V_H;3} \prod_{a=1}^{\mathrm{Dim} V_2} dS_a^{V_2} \prod_{a=1}^{\mathrm{Dim} V_3} dS_a^{V_3}$$

We are now able to apply standard techniques from Gaussian integration to calculate expectation values of observables

$$\langle f(M)\rangle = \frac{1}{\mathcal{Z}}\int dM e^{-\mathcal{S}} f(M)$$

Which we can generate by taking derivatives of the result

$$\mathcal{Z} = \int dM \exp\left(-\frac{1}{2}x_\alpha \Lambda_{\alpha\beta} x_\beta + \rho_\alpha x_\alpha\right) = \sqrt{\frac{(2\pi)^N}{\det\Lambda}} \exp\left(\frac{1}{2}\rho_\alpha (\Lambda^{-1})_{\alpha\beta}\rho_\beta\right)$$

The only non-vanishing first order expectation values are

$$\langle S^{V_0;\alpha} \rangle = \sum_{\beta} (\Lambda_{V_0}^{-1})_{\alpha\beta} \rho_{\beta}^{V_0}$$

From this we can write the first order expectation values of the original M_{ij}

$$\begin{aligned} \langle M_{ij} \rangle &= \frac{1}{D} \langle S^{00} \rangle + \frac{1}{\sqrt{D-1}} \sum_{a=1}^{D-1} C_{a,i} C_{a,j} \langle S^{HH\to 0} \rangle \\ &= \frac{(\Lambda_{V_0}^{-1})_{1\beta} \rho_{\beta}^{V_0}}{D} + \frac{(\Lambda_{V_0}^{-1})_{2\beta} \rho_{\beta}^{V_0}}{\sqrt{D-1}} F(i,j) \end{aligned}$$

The quadratic expectation values are given by

$$\begin{aligned}\langle S_a^{V_i;\alpha} S_b^{V_j;\beta} \rangle_{\text{conn}} &\equiv \langle S_a^{V_i;\alpha} S_b^{V_j;\beta} \rangle - \langle S_a^{V_i;\alpha} \rangle \langle S_b^{V_j;\beta} \rangle \\ &= \delta(V_i, V_j)(\Lambda_{V_i}^{-1})_{\alpha\beta}\delta_{ab}\end{aligned}$$

The decoupling of the different irreps follows from the factorised form of the partition function

Since we have a Gaussian theory higher point expectation values can be written in terms of linear and quadratic expectation values by Wick's theorem

$$\begin{aligned}\langle M_{ij}M_{kl}M_{pq}\rangle =&\langle M_{ij}M_{kl}\rangle_{\text{conn}}\langle M_{pq}\rangle + \langle M_{ij}M_{pq}\rangle_{\text{conn}}\langle M_{kl}\rangle\\ &+ \langle M_{kl}M_{pq}\rangle_{\text{conn}}\langle M_{ij}\rangle + \langle M_{ij}\rangle\langle M_{kl}\rangle\langle M_{pq}\rangle\end{aligned}$$

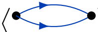
where $\langle M_{ij}M_{kl}\rangle_{\text{conn}}$ is given by the following expression

Permutation invariant Gaussian matrix model

Quadratic expectation values

$$
\begin{aligned}
\langle M_{ij} M_{kl} \rangle_{\text{conn}} &= \frac{1}{D^2} (\Lambda_{V_0}^{-1})_{11} + \frac{(\Lambda_{V_0}^{-1})_{22}}{D-1} F(i,j) F(k,l) + \frac{(\Lambda_{V_0}^{-1})_{12}}{D\sqrt{D-1}} \Big(F(k,l) + F(i,j) \Big) \\
&+ \frac{(\Lambda_{V_H}^{-1})_{11}}{D} F(j,l) + \frac{(\Lambda_{V_H}^{-1})_{22}}{D} F(i,k) + \frac{D(\Lambda_{V_H}^{-1})_{33}}{(D-2)} \sum_{p,q=1}^{D} F(i,p) F(j,p) F(k,q) F(l,q) F(p,q) \\
&+ \frac{(\Lambda_{V_H}^{-1})_{12}}{D} \Big(F(j,k) + F(i,l) \Big) + \frac{(\Lambda_{V_H}^{-1})_{13}}{\sqrt{D-2}} \sum_{p=1}^{D} \Big(F(j,p) F(k,p) F(l,p) + F(i,p) F(j,p) F(l,p) \Big) \\
&+ \frac{(\Lambda_{V_H}^{-1})_{23}}{\sqrt{D-2}} \sum_{p=1}^{D} \Big(F(i,p) F(k,p) F(l,p) + F(i,p) F(j,p) F(k,p) \Big) \\
&+ (\Lambda_{V_2}^{-1}) \bigg(\frac{1}{2} F(i,k) F(j,l) + \frac{1}{2} F(i,l) F(j,k) - \frac{D}{D-2} \sum_{p,q=1}^{D} F(i,p) F(j,p) F(k,q) F(l,q) F(p,q) \\
&- \frac{1}{(D-1)} F(i,j) F(k,l) \bigg) + \frac{(\Lambda_{V_3}^{-1})}{2} \Big(F(i,k) F(j,l) - F(i,l) F(j,k) \Big).
\end{aligned}
$$

As as example

$\langle$  $\rangle = \sum_{i,j=1}^{D} \langle M_{ij} M_{ij} \rangle_{\text{conn}}$

$$
\begin{aligned}
&= \big((\Lambda_{V_0}^{-1})_{1\alpha}\rho_{\alpha}^{V_0}\big)^2 + \big((\Lambda_{V_0}^{-1})_{2\alpha}\rho_{\alpha}^{V_0}\big)^2 \\
&+ (\Lambda_{V_0}^{-1})_{11} + (\Lambda_{V_0}^{-1})_{22} \\
&+ (D-1)\big((\Lambda_{V_H}^{-1})_{11} + (\Lambda_{V_H}^{-1})_{22} + (\Lambda_{V_H}^{-1})_{33}\big) \\
&+ \frac{D(D-3)}{2}(\Lambda_{V_2}^{-1}) \\
&+ \frac{(D-1)(D-2)}{2}(\Lambda_{V_3}^{-1})
\end{aligned}
$$

Experiment

This model was compared to experimental data in order to test Gaussianity (Ramgoolam, Sadrzadeh, Sword '19)

A data set of adjective and intransitive verb matrices was constructed from large corpora of data

The couplings of the theoretical model were set by evaluating the first and second order expectation values of observables

This was then used to predict a selection of cubic and quartic expectation values which could be compared with experiment

Graph	Expectation value	Theoretical val.	Experimental val.	Ratio
1	$\sum_i \langle (M_{ii})^3 \rangle$	1.44×10^{-1}	2.52×10^{-1}	0.57
2	$\sum_{i,j} \langle (M_{ij})^3 \rangle$	8.43×10^{-1}	3.65	0.23
3	$\sum_{i,j,k} \langle M_{ij} M_{jk} M_{ki} \rangle$	1.68	10.6	0.16
4	$\sum_{i,j,k} \langle M_{ij} M_{jj} M_{jk} \rangle$	53.8	80.1	0.67
5	$\sum_{i,j,k,l} \langle M_{ij} M_{kk} M_{ll} \rangle$	2.94×10^6	3.03×10^6	0.97
6	$\sum_{i,j,k,l} \langle M_{ij} M_{jk} M_{ll} \rangle$	4.83×10^4	5.04×10^4	0.96
7	$\sum_{i,j,k,l,m} \langle M_{ij} M_{kl} M_{mm} \rangle$	5.93×10^7	6.01×10^7	0.99
8	$\sum_{i,j,k,l,m,n} \langle M_{ij} M_{kl} M_{mn} \rangle$	1.38×10^9	1.40×10^9	0.98
9	$\sum_{i_1 \ldots i_7} \langle M_{i_1 i_2} M_{i_3 i_4} M_{i_5 i_6} M_{i_7 i_7} \rangle$	7.83×10^{10}	8.14×10^{10}	0.96
10	$\sum_{i_1 \ldots i_8} \langle M_{i_1 i_2} M_{i_3 i_4} M_{i_5 i_6} M_{i_7 i_8} \rangle$	1.86×10^{12}	1.96×10^{12}	0.95

Table 1: Comparison of theoretical and experimental expectation values for adjectives

Graph	Expectation value	Theoretical val.	Experimental val.	Ratio
1	$\sum_i \langle (M_{ii})^3 \rangle$	1.76×10^{-1}	3.22×10^{-1}	0.55
2	$\sum_{i,j} \langle (M_{ij})^3 \rangle$	9.36×10^{-1}	4.26	0.22
3	$\sum_{i,j,k} \langle M_{ij} M_{jk} M_{ki} \rangle$	1.62	9.98	0.16
4	$\sum_{i,j,k} \langle M_{ij} M_{jj} M_{jk} \rangle$	51.2	73.7	0.70
5	$\sum_{i,j,k,l} \langle M_{ij} M_{kk} M_{ll} \rangle$	2.87×10^6	2.92×10^6	0.99
6	$\sum_{i,j,k,l} \langle M_{ij} M_{jk} M_{ll} \rangle$	4.12×10^4	4.32×10^4	0.95
7	$\sum_{i,j,k,l,m} \langle M_{ij} M_{kl} M_{mm} \rangle$	5.32×10^7	5.35×10^7	0.99
8	$\sum_{i,j,k,l,m,n} \langle M_{ij} M_{kl} M_{mn} \rangle$	1.20×10^9	1.26×10^9	0.95
9	$\sum_{i_1 \dots i_7} \langle M_{i_1 i_2} M_{i_3 i_4} M_{i_5 i_6} M_{i_7 i_7} \rangle$	6.97×10^{10}	7.27×10^{10}	0.96
10	$\sum_{i_1 \dots i_8} \langle M_{i_1 i_2} M_{i_3 i_4} M_{i_5 i_6} M_{i_7 i_8} \rangle$	1.66×10^{12}	1.85×10^{12}	0.90

Table 2: Comparison of theoretical and experimental expectation values for verbs

Generally a good agreement, between 90% and 99% , even in the worst case the theory gives the correct order of magnitude.

These results improve for lower values of D

The authors note that this suggests the discrepancy may be related to the specifics of the method used to construct the verb and adjective matrices

Extending the model

Taking inspiration from linguistics lead us to consider permutation invariant Gaussian one-matrix models

$$(red)_{ij}(box)_j = (red\ box)_i$$

Could also consider:

Adjective verb combinations: Two-matrix models of M_{ij} and N_{ij}

$$(eat)_{ij}(quickly)_{jk}(cats)_k = (cats\ eat\ quickly)_i$$

Transitive verbs: Tensors models of T_{ijk}

$$(like)_{ijk}(cats)_j(fish)_k = (cats\ like\ fish)_i$$

Can apply similar techniques to solve permutation invariant Gaussian two-matrix models (GB, Padellaro, Ramgoolam '21)

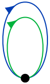
The observables of the theory $f(M, N)$ are again permutation invariant polynomials and must satisfy

$$f(M_{i,j}, N_{k,l}) = f(M_{\sigma(i),\sigma(j)} N_{\sigma(k),\sigma(l)}), \qquad \forall \sigma \in S_D$$

There is a 1-1 correspondence between these observables and directed, coloured graphs

Extensions

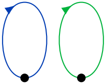
Two matrix model: Observable graph correspondence



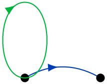
$$\sum_{i} M_{ii}N_{ii}$$

$$\sum_{i,j} M_{ij}N_{ij}$$

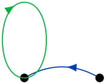
$$\sum_{i,j} M_{ji}N_{ij}$$



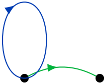
$$\sum_{i,j} M_{ii}N_{jj}$$



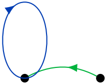
$$\sum_{i,j} M_{ij}N_{ii}$$



$$\sum_{i,j} M_{ji}N_{ii}$$



$$\sum_{i,j} M_{ii}N_{ij}$$



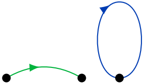
$$\sum_{i,j} M_{ii}N_{ji}$$

$$\sum_{i,j,k} M_{jk}N_{ij}$$

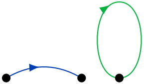
$$\sum_{i,j,k} M_{ij}N_{ik}$$

$$\sum_{i,j,k} M_{kj}N_{ij}$$

$$\sum_{i,j,k} M_{ij}N_{jk}$$



$$\sum_{i,j,k} M_{kk}N_{ij}$$



$$\sum_{i,j,k} M_{ij}N_{kk}$$

$$\sum_{i,j,k,l} M_{ij}N_{kl}$$

Summary

- A large collection of matrices that arises in computational linguistics are permutation invariant and Gaussian
- We considered a permutation invariant Gaussian one-matrix model to study the statistics of these matrices
- The observables of the theory are in 1-1 correspondence with directed graphs
- With the help of representation theory of the symmetric group this model is solvable
- The theoretical expectation value predictions are mostly in good agreement with experimental values

Summary

- More recently, the most general permutation invariant Gaussian two-matrix model was solved
- A direction for future work is to consider equivalent tensor models, T_{ijk}