Why is there only one definition of eigenvalue?

Henry Towsner

University of Pennsylvania

February 29, 2020

Observation

There's only one notion of an eigenvalue of a matrix.

Observation

There's only one notion of an eigenvalue of a matrix.

This is directly related to the fact that there is only one notion of graph quasirandomness.

Compare what happens to matrices with what happens to *tensors*. For our purposes, it will suffice to consider 3-*hypermatrices*—that is, three dimensional arrays of numbers

$$T = [t_{ijk}].$$

Compare what happens to matrices with what happens to *tensors*. For our purposes, it will suffice to consider 3-*hypermatrices*—that is, three dimensional arrays of numbers

$$T = [t_{ijk}].$$

We usually define the eigenvalues of a matrix M to be the value λ so there is a vector v such that $Mv = \lambda v$. But this defintion doesn't have a clear analog for a tensor: if v is a vector, Tv is a matrix, while if M is a matrix, $T \cdot M$ is a vector.

But there is an alternative definition of the eigenvalue of a matrix: an eigenvalue is a local maximum of the inner product

$$\lambda = \langle Mv, v \rangle = \sum_{i,j} m_{ij} v_i v_j$$

among the unit vectors v.

But there is an alternative definition of the eigenvalue of a matrix: an eigenvalue is a local maximum of the inner product

$$\lambda = \langle Mv, v \rangle = \sum_{i,j} m_{ij} v_i v_j$$

among the unit vectors v.

This definition does make sense: we can define a *vector eigenvalue* of T to be a local maximum of

$$\lambda = \sum_{i,j,k} t_{ijk} v_i v_j v_k$$

over unit vectors v.

With symmetric matrices, the second eigenvalue measures deviation from quasirandomness.

With symmetric matrices, the second eigenvalue measures deviation from quasirandomness.

That is, suppose M is a large, square, symmetric, dense matrix:

- M is an $n \times n$ matrix with n big,
- $m_{ij} = m_{ji}$,
- the entries $m_{ij} \in [0, 1]$,
- the average value $\sum_{i,j} m_{ij} = pn^2$.

Then the first eigenvalue measures the even distribution of M:

Lemma

- $\lambda_1 \geq np$, *and*
- $\lambda_1 = np$ *if and only if, for each* $i \leq n$, $\sum_j m_{ij} = np$.

Then the first eigenvalue measures the even distribution of M:

Lemma

- $\lambda_1 \geq np$, *and*
- $\lambda_1 = np$ *if and only if, for each* $i \leq n$, $\sum_j m_{ij} = np$.

Hint: consider the unit vector constantly equal to $1/\sqrt{n}$.

Then the first eigenvalue measures the even distribution of M:

Lemma

- $\lambda_1 \geq np$, *and*
- $\lambda_1 = np$ *if and only if, for each* $i \leq n$, $\sum_j m_{ij} = np$.

Hint: consider the unit vector constantly equal to $1/\sqrt{n}$.

Suppose $\lambda_1 \approx np$. We can then look at λ_2.

Suppose we generate M randomly: for each ij, we flip a fair coin independently to decide whether m_{ij} is 0 or 1. Then, with high probability, $\lambda_1 \approx \frac{n}{2}$ and λ_2 is close to 0 (and therefore all other eigenvalues are also close to 0).

One way to see this: it is a standard fact about eigenvalues that

$$\sum_i \lambda_i^4 = \mathrm{tr}(M^4)$$

One way to see this: it is a standard fact about eigenvalues that

$$\sum_i \lambda_i^4 = \mathrm{tr}(M^4) = \sum_{ijkl} m_{ij} m_{jk} m_{kl} m_{li}.$$

For most $ijkl$, these are four different pairs, so each quadruple has a $1/2^4$ chance of being present, so this sum is $\frac{1}{2^4} n^4$. If we think of M as the adjacency matrix of a graph, this amounts to counting the number of cycles of length 4.

One way to see this: it is a standard fact about eigenvalues that

$$\sum_i \lambda_i^4 = \mathrm{tr}(M^4) = \sum_{ijkl} m_{ij} m_{jk} m_{kl} m_{li}.$$

For most $ijkl$, these are four different pairs, so each quadruple has a $1/2^4$ chance of being present, so this sum is $\frac{1}{2^4} n^4$. If we think of M as the adjacency matrix of a graph, this amounts to counting the number of cycles of length 4.

Since $\lambda_1 \approx \frac{n}{2}$, there is no room left for the other eigenvalues: $\sum_{1<i\leq n} \lambda_i^4$ must be small.

Furthermore, this is, suitably interpreted, an equivalence:

Theorem

There is a language $\mathcal{L}$ for probabilistic matrices such that, for each sentence σ and each $p \in (0,1)$, if σ is (with high probability) true in the random matrix with density p then there is an ϵ so that σ holds in any matrix with n large enough, $\lambda_1 - np < \epsilon$, and $|\lambda_2| < \epsilon$.

For tensors, this doesn't work. λ_1 still measures even distribution:

Lemma

- $\lambda_1 \geq n^{3/2}p$, *and*
- $\lambda_1 = n^{3/2}p$ *if and only if, for each* $i \leq n$, $\sum_{j,k} a_{ijk} = n^2 p$

For tensors, this doesn't work. λ_1 still measures even distribution:

Lemma

- $\lambda_1 \geq n^{3/2}p$, *and*
- $\lambda_1 = n^{3/2}p$ *if and only if, for each* $i \leq n$, $\sum_{j,k} a_{ijk} = n^2 p$

but consider the tensor generated as follows:

- Generate a random *matrix* $B = [b_{ij}]$ by flipping a coin for each pair i, j.
- Set $t_{ijk} = b_{ij} + b_{ik} + b_{jk} \mod 2$.

Then T is evenly distributed and has a small λ_2, but is very different from a random tensor.

For tensors, this doesn't work. λ_1 still measures even distribution:

Lemma

- $\lambda_1 \geq n^{3/2} p$, *and*
- $\lambda_1 = n^{3/2} p$ *if and only if, for each* $i \leq n$, $\sum_{j,k} a_{ijk} = n^2 p$

but consider the tensor generated as follows:

- Generate a random *matrix* $B = [b_{ij}]$ by flipping a coin for each pair i, j.
- Set $t_{ijk} = b_{ij} + b_{ik} + b_{jk} \mod 2$.

Then T is evenly distributed and has a small λ_2, but is very different from a random tensor. For instance, in a random tensor we would have

$$\frac{1}{n^4} \sum_{ijkl} r_{ijk} r_{ijl} r_{ikl} r_{jkl} \approx 1/2^4$$

while

$$\frac{1}{n^4} \sum_{ijkl} t_{ijk} t_{ijl} t_{ikl} t_{jkl} \approx 1/2^3.$$

We could have defined a different kind of eigenvalue:

Definition

When T is a symmetric tensor, a *matrix eigenvalue of* T is a value λ which is a local maximum of

$$\sum_{i,j,k} t_{ijk} m_{ij} m_{ik} m_{jk}$$

among unit matrices M. We call such an M an *eigenmatrix* of T.

The first eigenvalue measures a stronger kind of even distribution:

Lemma

- $\lambda_1 \geq n^{-3}p$,
- $\lambda_1 = n^{-3}p$ *if and only if, for every* i, j, $\sum_k a_{ijk} = np$.

The first eigenvalue measures a stronger kind of even distribution:

Lemma

- $\lambda_1 \geq n^{-3}p$,
- $\lambda_1 = n^{-3}p$ *if and only if, for every* i, j, $\sum_k a_{ijk} = np$.

The second *matrix* eigenvalue does actually measure deviation from quasirandomness.

But there are other notions:

- eigenvalues are local maxima of $\sum_{i,j,k} t_{ijk} m_{ij} v_k$ among unit matrices M and unit vectors v (this is the notion related to Tao's *slice rank*),

But there are other notions:

- eigenvalues are local maxima of $\sum_{i,j,k} t_{ijk} m_{ij} v_k$ among unit matrices M and unit vectors v (this is the notion related to Tao's *slice rank*),
- eigenvalues are local maxima of $\sum_{i,j,k} t_{ijk} m_{ij} m_{ik}$ among unit matrices M

But there are other notions:

- eigenvalues are local maxima of $\sum_{i,j,k} t_{ijk} m_{ij} v_k$ among unit matrices M and unit vectors v (this is the notion related to Tao's *slice rank*),
- eigenvalues are local maxima of $\sum_{i,j,k} t_{ijk} m_{ij} m_{ik}$ among unit matrices M
- etc.

Suppose we have a d-tensor. There is a notion of eigenvalue corresponding to every antichain of proper subsets of $\{1, \ldots, d\}$, identifying antichains under permutations of $\{1, \ldots, d\}$.

Suppose we have a d-tensor. There is a notion of eigenvalue corresponding to every antichain of proper subsets of $\{1, \ldots, d\}$, identifying antichains under permutations of $\{1, \ldots, d\}$.

So for 3-tensors, we have antichains like:

- $\{\{1\}, \{2\}, \{3\}\}$ corresponding to vector eigenvalues,
- $\{\{1, 2\}, \{1, 3\}, \{2, 3\}\}$ corresponding to matrix eigenvalues,
- $\{\{1, 2\}, \{3\}\}$,
- $\{\{1, 2\}, \{1, 3\}\}$,
- etc.

Question

How can we be sure we've found them all?

Question

How can we be sure we've found them all?

Notions of eigenvalues correspond to "canonical" sub-σ-algebras in a certain setting, but it's not clear how to pin down the notion of canonical.

Question

How can we be sure we've found them all?

Notions of eigenvalues correspond to "canonical" sub-σ-algebras in a certain setting, but it's not clear how to pin down the notion of canonical.

The end.