2/25

Schur functions: the "most important" basis of Sym

We now define the Schur functions $s_\lambda(x_1, x_2, \dots)$, which are another basis of Sym - and (arguably) the most important one.

It's a bit hard to motivate what makes them so important:

i) W.r.t. the inner product $\langle \cdot, \cdot \rangle : \mathrm{Sym} \times \mathrm{Sym} \to \mathbb{R}$ just mentioned they are orthonormal: $\langle s_\lambda, s_\mu \rangle = \begin{cases} 1 & \text{if } \lambda = \mu, \\ 0 & \text{otherwise} \end{cases}$

ii) In the representation theory interpretations of Sym, they correspond to "irreducible representations."

The definition of $s_\lambda(x_1, x_2, \dots)$ will be very different from other bases:

DEF'N Let $\lambda \vdash n$ $= (\lambda_1, \lambda_2, \dots, \lambda_k)$ be a partition. Recall that [redacted] λ's Young diagram has λ_i boxes in the i^{th} row:

$\lambda = (3,3,1) \Leftrightarrow \lambda =$  (rows are left-justified)

A semistandard Young tableau of shape λ is a filling of the boxes of its Young diagram w/ positive integers such that:

- entries are weakly increasing along rows
- entries are strictly increasing down columns

e.g. $T =$

1	1	2
2	2	3
4		

($\leq$ along rows, $\uparrow$ strictly down columns) is a SSYT of shape (3,3,1)

For T a SSYT, content (T) is vector $co(T) := (c_1, c_2, \dots) \in \mathbb{N}^\infty$ where $c_i = \#$ boxes w/ entry i

e.g. $co(T) = (2, 3, 1, 1, 0, 0, \dots)$

The Schur function $s_\lambda(x_1, x_2, \dots)$ is then

$$s_\lambda := \sum_{\substack{T \text{ SSYT} \\ \text{of sh.} = \lambda}} \vec{x}^{\,co(T)} = \sum_{\substack{T \text{ SSYT} \\ sh(T) = \lambda}} \prod_{i \geq 1} x_i^{c_i(T)} \in \mathbb{C}[[x_1, x_2, \dots]]$$

__e.g.__ Let $\lambda = (2,1)$. Let's compute the __Schur polynomial__

$$s_\lambda(x_1, x_2, x_3) = s_\lambda(x_1, x_2, x_3, 0, 0, \dots)$$

The SSYT of sh. $= (2,1)$ and entries in $\{1,2,3\}$ are:

$$\begin{array}{cc}1&1\\2&\end{array}\quad \begin{array}{cc}1&1\\3&\end{array}\quad \begin{array}{cc}1&2\\2&\end{array}\quad \begin{array}{cc}1&2\\3&\end{array}\quad \begin{array}{cc}2&2\\3&\end{array}\quad \begin{array}{cc}2&3\\3&\end{array}\quad \begin{array}{cc}1&3\\2&\end{array}\quad \begin{array}{cc}1&3\\3&\end{array}$$

So

$$s_\lambda(x_1,x_2,x_3) = x_1^2x_2 + x_1^2x_3 + x_2^2x_1 + x_2^2x_3 + x_3^2x_1 + x_3^2x_2 + 2x_1x_2x_3$$

$$= 2m_{(1,1,1)}(x_1,x_2,x_3) + m_{(2,1)}(x_1,x_2,x_3)$$

a __Symmetric polynomial__! ⟵ not a priori obvious it should be symmetric

__In fact__, $s_{(2,1)}(x_1, x_2, \dots) = 2m_{(1,1,1)} + m_{(2,1)} \in \mathrm{Sym}$.

Schur functions generalize elementary + complete homo. sym. fn's:

__Prop__:
- $s_{(1^n)}(x_1, \dots) = e_n(x_1, \dots)$
- $s_{(n)}(x_1, \dots) = h_n(x_1, \dots)$

__Pf__:

$$s_{(1^n)}(x_1, \dots) = s_{\text{column}}(x_1, \dots) = \sum_{T \text{ SSYT sh} = \text{column}} \vec{x}^{\,co(T)}$$

But an SSYT of sh. = column is just $\begin{array}{c}i_1\\i_2\\\vdots\\i_n\end{array}$ w/ $i_1 < i_2 < \cdots < i_n$

So indeed $s_{(1^n)} = \sum_{i_1 < \cdots < i_n} x_{i_1} \cdots x_{i_n} = e_n$.

Similarly $s_{(n)} = \sum_{T \text{ SSYT sh.} = \text{row}} \vec{x}^{\,co(T)}$ and an SSYT of sh. row is $\begin{array}{cccc}i_1&i_2&\cdots&i_n\end{array}$ w/ $i_1 \le \cdots \le i_n$

So $s_{(n)} = \sum_{i_1 \le \cdots \le i_n} x_{i_1} \cdots x_{i_n} = h_n$. ∎

2/28

But for other shapes than single row/column, not clear that s_λ is symmetric. We prove this now.

DEF'N Let T be a SSYT. The i^{th} Bender-Knuth involution (for $i=1,2,\dots$) applied to T, denoted $b_i(T)$, is the following operation:

- first, "freeze" all entries i ~~immediately~~ (directly) above an $i+1$, and all $i+1$'s (directly) below an i,
- then, in each row, if there are a unfrozen i's and b unfrozen $i+1$'s in this row, modify these entries so that there are b unfrozen i's and a unfrozen $i+1$'s (in unique way that preserves SSYT-ness).

e.g. Let's apply b_4 to

$T =$

1	1	1	1	2	2	2	3	3	3	4
2	2	3	3	3	3	4	4	4	5	
3	4	4	4	5	5	5	5			
5	5	5	6							
6	6									

$\square$ = frozen

← 1 unfrozen 4, 2 unfrozen 5's in 3rd row

$b_4(T) =$

1	1	1	1	2	2	2	3	3	3	5
2	2	3	3	3	3	4	4	4	5	
3	4	4	4	4	5	5	5			
4	5	5	6							
6	6									

← 2 unfrozen 4's, 1 unfrozen 5 in 3rd row, etc...

Prop. $b_i(T)$ is an SSYT (of same shape as T!) with $co(b_i(T)) = (i, i+1)\cdot co(T)$, i.e., # i's in T = # $i+1$'s in $b_i(T)$ and vice-versa.
Also, $b_i(b_i(T)) = T$.

Pf: All statements are relatively straightforward. To see $co(b_i(T)) = (i,i+1)\cdot co(T)$, note that frozen i's + $i+1$'s come in pairs that cancel, while unfrozens get switched. ☒

Cor For any λ, s_λ is a symmetric function.

Pf: Bender-Knuth involutions show that $(i,i+1)\cdot s_\lambda = s_\lambda$

(since $\sum_{T \text{ SSYT } sh(T)=\lambda} \vec{x}^{co(T)} = \sum_{T: sh(T)=\lambda} \vec{x}^{co(b_i(T))} = \sum_{T: sh(T)=\lambda} \vec{x}^{(i,i+1)\cdot co(T)} = (i,i+1)\cdot \sum_{T: sh(T)=\lambda} \vec{x}^{co(T)}$.)

But then note that any permutation $\sigma \in S_n$ is a composition of adjacent transpositions $\sigma = (i_1, i_1+1)\cdot(i_2, i_2+1)\cdots(i_\ell, i_\ell+1)$

(Think about sorting numbers in a line: $\underline{7}\,\underline{1}\,\underline{3}\,\underline{2}\,\underline{5}\,\underline{6}\,\underline{4}$, can always do it by swapping adjacent positions.)

So $\sigma \cdot s_\lambda = s_\lambda$ for any $\sigma \in S_n$, so s_λ is symmetric! ☐

Thm $\{s_\lambda : \lambda \vdash n\}$ is a basis of $Sym(n)$.

Pf: Just proved that s_λ for $\lambda \vdash n$ is symmetric, and that it has degree n is clear. Since there are correct # of s_λ for a basis, what we need to show is that they span all of $Sym(n)$.

We do this, like with the other bases, by a triangularity argument. So write

$$s_\lambda = \sum_\mu K_{\lambda,\mu}\, m_\mu.$$

Note that $K_{\lambda,\mu} := \#$ SSYT w/ sh $=\lambda$ and co $=\mu$.

We claim that $K_{\lambda,\mu} \neq 0 \Rightarrow \mu \leq \lambda$ in lex. order.

Indeed, there is 1 tableau counted by $K_{\lambda,\lambda} = 1$: we have all i's in the i^{th} row

e.g.,

1	1	1	1
2	2	2	
3	3	3	
4			

Now suppose $\mu \neq \lambda$ and $K_{\lambda,\mu} \neq 0$. Let j be smallest # s.t. $\mu_j \neq \lambda_j$. Then $\lambda_i = \mu_i \ \forall i < j$

So a tableau counted by $K_{\lambda,\mu}$ has all i's in row i for $i<j$. So j^{th} row has $<\lambda_j$ j's $\Rightarrow \mu < \lambda$. ✓ ☐

3/2

<u>Expanding Schur functions in the other bases</u>

from what we just explained, we have

$$s_\lambda = \sum_\mu K_{\lambda,\mu}\, m_\mu$$

where $K_{\lambda\mu} := \#\{SSYT\ T : sh(T)=\lambda,\ co(T)=\mu\}$
← called the <u>Kostka numbers</u>

But we also have the e_μ, h_μ, and p_μ bases, so can ask what s_λ looks like in these bases.

The expansion of Schurs into power sums:

$$s_\lambda = \sum_\mu z_\mu^{-1} \chi^\lambda(\mu)\, p_\mu \quad ; \quad z_\mu = 1^{m_1} m_1!\, 2^{m_2} m_2! \cdots \text{ if } \mu = 1^{m_1} 2^{m_2} \cdots$$

is perhaps the most important one, because

$\chi^\lambda(\mu)$ = "character of irreducible representation of S_n indexed by $\lambda \vdash n$ at permutation of cycle type μ."

There is a combinatorial formula for $\chi^\lambda(\mu)$ called the "<u>Murnaghan-Nakayama rule</u>"; see Ch. 7 Stanley EC2

But it's beyond scope of this class...

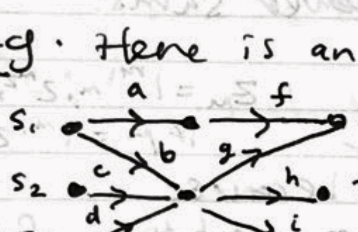
Instead we'll focus on the expansion of s_λ into e_μ/h_μ.

The formula for writing s_λ in the e_μ's/h_μ's is called the <u>Jacobi-Trudi formula</u> and it expresses s_λ as a <u>determinant</u>.

To prove the J-T formula, we will apply another result: the <u>Lindström-Gessel-Viennot lemma</u> which is itself a very powerful enumerative tool worth knowing about.

DEF'N A *directed graph* (or *digraph*) $G=(V,E)$ has *vertex set* V and *directed edge set* E, where a directed edge $e=(u,v)$ is an ordered pair of vertices we draw as an arrow: $u \to v$.
We say G is *acyclic* if it has no *directed cycles*: $\circlearrowright$. An (acyclic) *network* is an acyclic digraph w/ distinguished *source vertices* $s_1, s_2, \ldots, s_n$ and *target vertices* $t_1, \ldots, t_n$, and a *weight function* $wt: E \to \mathbb{R}$ on edges.

e.g. Here is an acyclic network w/ 3 sources + targets:



w/ $a,b,c,d,e,f,g,h,i,j \in \mathbb{R}$ as edge weights.

DEF'N A *path* P in a digraph is a sequence of edges $s \xrightarrow{e_1} \bullet \xrightarrow{e_2} \cdots \bullet \xrightarrow{e_n} t$ connecting s to t.
We define the *weight* of P to be $wt(e_1) \cdot \cdots \cdot wt(e_n)$.

The *path matrix* of network G is $n \times n$ matrix M with $M_{i,j} := \sum_{\substack{\text{paths} \\ P: s_i \text{ to } t_j}} wt(P)$.

To a tuple $T = (P_1, \ldots, P_n)$ of paths we associate weight $wt(T) = wt(P_1) \cdot \cdots \cdot wt(P_n)$. We say the tuple is *nonintersecting* if all the vertices in P_i and P_j are disjoint, for every $i \neq j$.

Thm (Lindström-Gessel-Viennot Lemma)

Let M be the path matrix of acyclic network G.

Then $$\det(M) = \sum_{\substack{\text{non-intersecting tuple} \\ T=(P_1,\dots,P_n) \\ P_i : s_i \to t_{\sigma(i)}}} \operatorname{sgn}(\sigma) \cdot \operatorname{wt}(T).$$

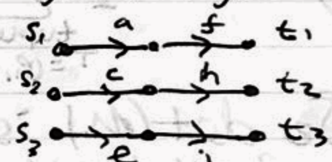
(Recall for a permutation $\sigma \in S_n$, $\operatorname{sgn}(\sigma) = (-1)^{\#\text{ inversions}(\sigma)}$.)

e.g. w/ G the network from previous example

path matrix is $$M = \begin{array}{c} \\ s_1 \\ s_2 \\ s_3 \end{array}\begin{bmatrix} t_1 & t_2 & t_3 \\ af+bg & bh & bi \\ cg & ch & ci \\ dg & dh & di+ej \end{bmatrix}$$

and $\det(M) = \big((af+bg)(ch)(di+ej) + (bh)(ci)(dg) + (bi)(cg)(dh)\big)$
$\quad - \big((bi)(ch)(dg) + (bh)(cg)(di+ej) + (af+bg)(ci)(dh)\big)$

$= (af)(ch)(ej)$ ← wt of unique tuple of non-int. lattice paths

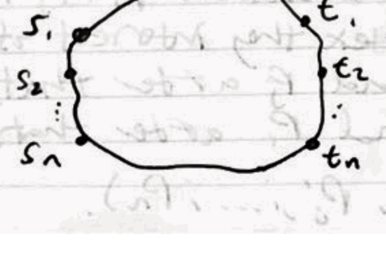


NOTE: $\operatorname{sgn}\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} = +1$ ✓

In this example, we see a very important special case:

Cor (Planar LGV lemma)

Suppose acyclic network G is planar (i.e., edges only cross at vertices) drawn in a disc w/ sources $s_1, \dots, s_n$ and targets $t_n, \dots, t_1$ on boundary (in counterclockwise order), like



Then, $$\det(M) = \sum_{\substack{\text{non-intersecting} \\ T=(P_1,\dots,P_n) \\ P_i : s_i \to t_i}} \operatorname{wt}(T).$$

//

3/2 #2

Pf of LGV Lemma: (from Ch. 2 of Sagan)

We will use a technique from last semester: Sign-reversing involution.

First, we use the "Leibniz formula" for determinant:

$$\det(M) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i \geq 1} M_{i,\sigma(i)}$$

$$= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i \geq 1} \sum_{\substack{\text{path} \\ P: s_i \to t_{\sigma(i)}}} wt(P)$$

$$= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \sum_{\substack{\text{tuple of paths} \\ T = (P_1, \dots, P_n) \\ P_i: s_i \to t_{\sigma(i)}}} wt(T)$$

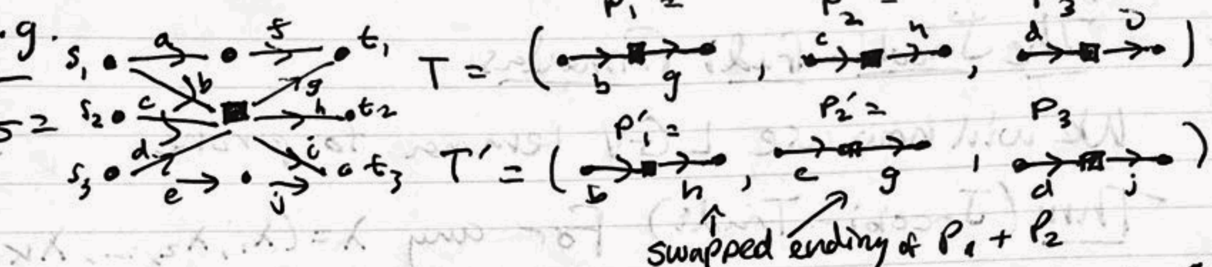
$$= \sum_{\substack{\text{tuple of paths} \\ T = (P_1, \dots, P_n),\ P_i: s_i \to t_{\sigma(i)}}} \operatorname{sgn}(\sigma) \cdot wt(T)$$

So $\det(M)$ is naturally the generating function of all tuples of paths connecting sources to sinks.

To show that this sum can be taken over only the non-intersecting tuples, we will cancel all the intersecting tuples, by defining an appropriate sign-reversing involution:

- given an intersecting tuple $T = (P_1, \dots, P_n)$, let (i,j) be lex. smallest pair such that $P_i + P_j$ intersect, and let v be the last vertex they intersect at. Define P_i' to be P_i up to v, and P_j after that, and P_j' to be P_j up to v, and P_i after that. Set $T' := (P_1, P_2, \dots, P_i', \dots, P_j', \dots, P_n)$.

e.g. G =



Then $T \mapsto T'$ is an involution, and if σ and σ' are the permutations corresponding to T, T' we have $sgn(\sigma) = -sgn(\sigma')$ because $\sigma' = \sigma \cdot (i,j)$ [Exercise: check this fact about permutation signs.]

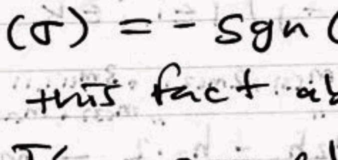
Also, T and T′ use same edges, so $wt(T) = wt(T')$.

Thus $T \mapsto T'$ is a sign-reversing involution on all intersecting tuples, so the intersecting tuples cancel in the sum and we get

$$\det(M) = \sum_{\substack{\text{tuples} \\ T=(P_1,\dots,P_n):\, P_i: s_i \to t_{\sigma(i)}}} sgn(\sigma) \cdot wt(T) = \sum_{\substack{\text{non-intersecting} \\ T=(P_1,\dots,P_n):\, P_i: s_i \to t_{\sigma(i)}}} sgn(\sigma) \cdot wt(T).$$ ☒

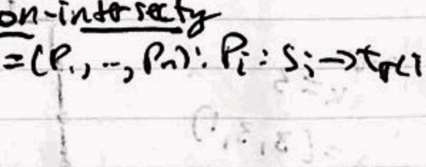
Pf of planar LGV corollary:

If G looks like



and is planar,

then for any $T = (P_1, \dots, P_n)$ whose σ is not $\begin{pmatrix} 1 & 2 & \cdots & n \\ 1 & 2 & \cdots & n \end{pmatrix}$ we will have an intersection: there will be some $i < j$ with $\sigma(i) > \sigma(j)$.



So for planar networks like this, we only need to sum over T's with σ = identity, which have $sgn(\sigma) = +1$. ☒

3/14

The Jacobi-Trudi Formulas

We will now use LGV lemma to prove.

Thm (Jacobi-Trudi) For any $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_k)$,

(a) $s_\lambda = \det(h_{\lambda_i - i + j})_{1 \le i,j \le k}$

(b) $s_\lambda = \det(e_{\lambda_i' - i + j})_{1 \le i,j \le k}$

e.g. $s_{(2,1)} = \det\begin{bmatrix} h_2 & h_3 \\ h_0 & h_1 \end{bmatrix} = h_2 h_1 - h_3 \cdot 1 = m_{(2,1)} + 2m_{(1,1,1)}$ ✓

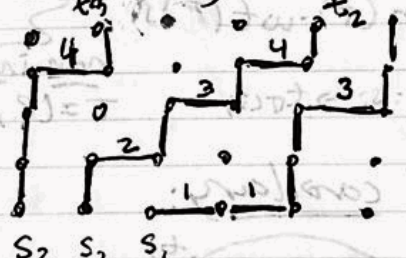
(where $h_2h_1 = m_{(3)} + 2m_{(2,1)} + 3m_{(1,1,1)}$ and $h_3 = m_{(3)} + m_{(2,1)} + m_{(1,1,1)}$)

NOTE $h_r = e_r = 0$ for $r < 0$ in this formula.

Pf of Jacobi-Trudi: We prove (a); (b) is similar.

Construct the following network G based on λ:

e.g. $k=3$, $\lambda = (3,3,1)$



$\longleftrightarrow$

1	1	3
2	3	4
4		

The network is a part of the grid $\mathbb{Z}^2$, w/ all edges directed right and up. Our sources are $s_i = (k-i, 1)$ and targets are $t_i = (k - i + \lambda_i, n)$ for $i = 1, 2, \dots, k$. (Here n is a number will will send $\to \infty$.)

As depicted above, tuples (P_1, ..., P_k) of non-intersecting lattice paths w/ $P_i : s_i \to t_i$ correspond bijectively to SSYT of shape $= \lambda$

3/16

(w/ entries $\le n$): the corresponding tableau T has entries $i_1, \dots, i_{\lambda_j}$ in the j^{th} row, where $i_1, \dots, i_{\lambda_j}$ are the horizontal step heights of path P_j.

There is something to check here: that the non-intersectingness of the paths is equivalent to the SSYT condition. That's an exercise for you...

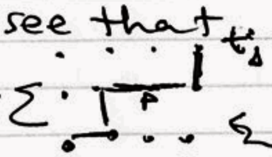
This bijection tells us what the edge weights of our network should be: vertical steps have wt $= 1$, and a horizontal step at height i has wt $= x_i$.

Thus the LGV Lemma applied to this network says

$$S_\lambda(x_1, \dots, x_n) = \det(M), \text{ where } M_{i,j} = \sum_{\substack{\text{paths} \\ p:\, s_i \to t_j}} \mathrm{wt}(p).$$

But it's not hard to see that

$$\sum_{\text{paths } P:\, s_i \to t_j} \mathrm{wt}(P) = \sum$$



$$= h_{\lambda_j - j + i}(x_1, \dots, x_n)$$

choose any size $\lambda_j - j + i$ multiset of heights of horiz. steps.

So $S_\lambda(x_1, \dots, x_n) = \det(h_{\lambda_j - j + i}) = \det(h_{\lambda_i - i + j})$, (transpose!)

and we get the J-T formula in limit $n \to \infty$. ☐

Rmk: Can use J-T formula (+ some determinant manipulation) to show

$$S_\lambda(x_1, \dots, x_n) = \frac{\det(x_j^{\lambda_i + n - i})}{\det(x_j^{n-i})}$$

"Vandermonde determinant" $= \prod_{1 \le j < i \le n} (x_j - x_i)$

which is actually the original definition of the Schur polynomials from late 19^{th}/early 20^{th} Century.

"Bialternant definition"