Computing Littlewood-Richardson Coefficients: Algorithms and Complexity

Hari Narayanan

MIT

Includes work with Ketan Mulmuley

Littlewood-Richardson coefficients

- Littlewood-Richardson coefficients are structure constants appearing in the representation theory of the general linear group (GL_n) and count toy analogs of obstructions in GCT.

Littlewood-Richardson coefficients

- Littlewood-Richardson coefficients are structure constants appearing in the representation theory of the general linear group (GL_n) and count toy analogs of obstructions in GCT.
- Ubiquitous in math (representation theory, algebraic combinatorics, tilings, ...)

Littlewood-Richardson coefficients

- Littlewood-Richardson coefficients are structure constants appearing in the representation theory of the general linear group (GL_n) and count toy analogs of obstructions in GCT.
- Ubiquitous in math (representation theory, algebraic combinatorics, tilings, ...)
- and physics (fine structure of atomic spectra, ...)

Outline

1. Strongly polynomial randomized approximation scheme (SPRAS) for generic Littlewood-Richardson (LR) coefficients.

Outline

1. Strongly polynomial randomized approximation scheme (SPRAS) for generic Littlewood-Richardson (LR) coefficients.
2. For a semisimple Lie algebra G, assuming a mathematical conjecture for which there is extensive experimental evidence, algorithm for testing the positivity of all generalized LR coefficients.

Group Representations

- $\rho : G \to GL(V)$ homomorphism, then V is a representation of G

Group Representations

- $\rho : G \to GL(V)$ homomorphism, then V is a representation of G
- If no non-trivial proper subspace of V is mapped onto itself by all $g \in G$, V is *irreducible*.

Tensor product

- Tensor product $V \otimes W$

Tensor product

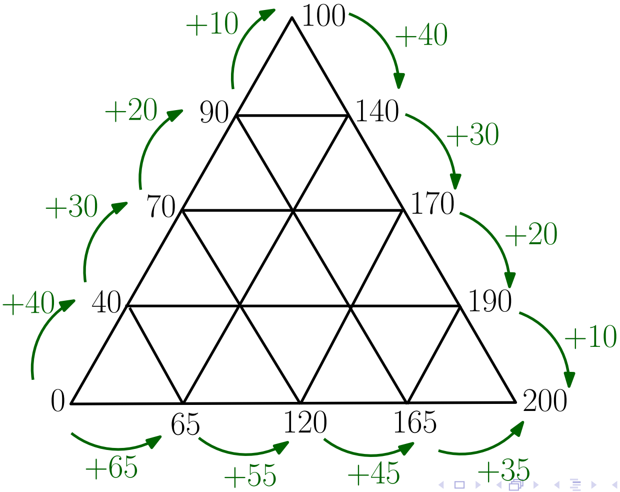
- Tensor product $V \otimes W$
  - $g(v \otimes w) = gv \otimes gw.$

Tensor product

- Tensor product $V \otimes W$
  - $g(v \otimes w) = gv \otimes gw.$
- For V_ν, V_μ and V_λ irreducible representations of GL_n, the Littlewood-Richardson coefficient is given by $V_\lambda \otimes V_\mu := \bigoplus_\nu c^\nu_{\lambda\mu} V_\nu$

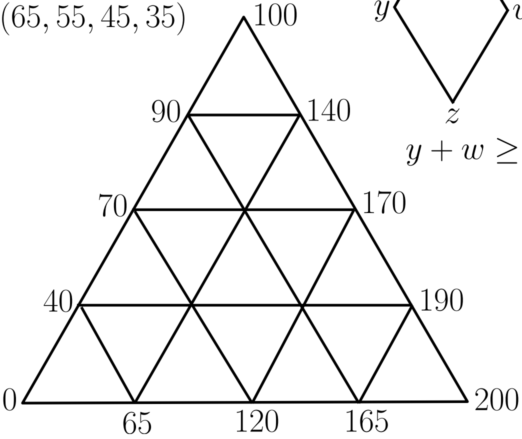
Hive model for LR coefficients

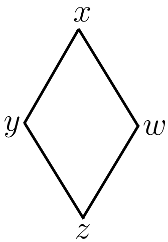
$$\lambda = (40, 30, 20, 10)$$
$$\mu = (40, 30, 20, 10)$$
$$\nu = (65, 55, 45, 35)$$



Hive model for LR coefficients

$$\lambda = (40, 30, 20, 10)$$
$$\mu = (40, 30, 20, 10)$$
$$\nu = (65, 55, 45, 35)$$





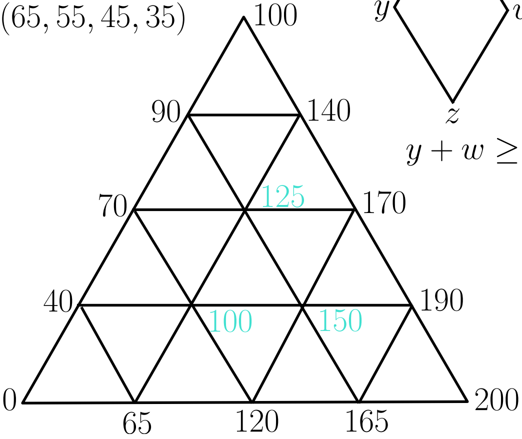
$$y + w \geq x + z$$

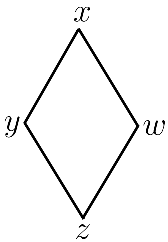
Hive model for LR coefficients

$\lambda = (40, 30, 20, 10)$

$\mu = (40, 30, 20, 10)$

$\nu = (65, 55, 45, 35)$





$y + w \geq x + z$

Honeycomb and Hive models

http://www.math.ucla.edu/~tao/java/Honeycomb.html

Complexity

Theorem (N'06)

LR coefficients for GL_n are $\#P-$complete.

Idea of Proof.

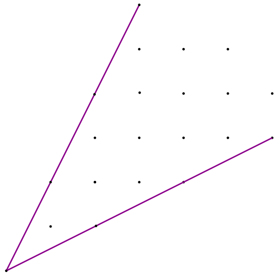
Reduce the computation of the number $I(a, b)$ of $2 \times n$ tables with given row and column sums a and b and positive integer entries (a known $\#P-$complete problem) to the computation of an LR coefficient $c^{\nu}_{\lambda\mu}$ via the Robinson-Schensted-Knuth correspondence and a way of compressing an exponential number of tableaux into a single tableau. □

LRcone

Let $eig(X)$ be the eigenvalues of X listed in non-increasing order.

LR cone $=$ cone of $3-$tuples $(eig(U), eig(V), eig(W))$ where

U, V, W are symmetric positive definite and $U + V = W$.



λ, μ and ν are partitions with n parts whose sizes monotonically decrease when $G = GL_n$.

$c^{\nu}_{\lambda\mu} > 0 \Leftrightarrow (\lambda, \mu, \nu)$ is a lattice point in the LR cone.

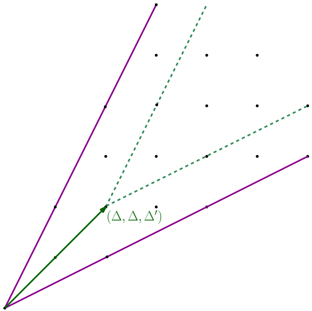
Approximating generic LR coefficients

Theorem (N'10)

There is a randomized strongly polynomial time algorithm for approximating a $1 - O(\gamma)$ *fraction of all* $c_{\lambda\mu}^{\nu}$ *corresponding to integer points in*

$$LRcone \cap \left\{ \|(\lambda, \mu, \nu)\|_1 \leq \frac{n^5}{\gamma} \right\}.$$

SPRAS for LR coefficients



$\alpha = \Omega(1/\log n)$,
$\Delta = \alpha(n^3, n^3 - n^2, \ldots, n^2)$,

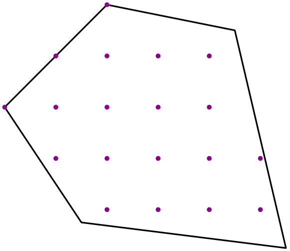
$\Delta' = \alpha(\frac{3n^3}{2} + \frac{n^2}{2}, \frac{3n^3}{2} - \frac{n^2}{2}, \ldots, \frac{n^3}{2} + \frac{3n^2}{2})$.

Theorem

If $(\lambda, \mu, \nu) \in (\Delta, \Delta, \Delta') + LRcone$, then $c^{\nu}_{\lambda\mu}$ can be approximated in randomized strongly polynomial time.

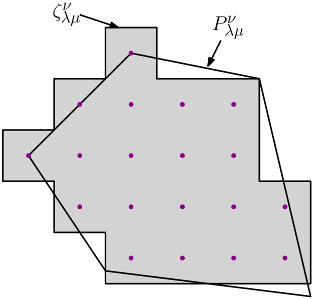
If $P_{\lambda\mu}^{\nu}$ is the hive polytope corresponding to (λ, μ, ν), the number of lattice points in $P_{\lambda\mu}^{\nu} \cap \mathbb{Z}^{\binom{n}{2}}$ is equal to the volume of

$$\zeta_{\lambda\mu}^{\nu} := \left\{ x \Big| \inf_{y \in P_{\lambda\mu}^{\nu} \cap \mathbb{Z}^{\binom{n}{2}}} \|x - y\|_\infty < \frac{1}{2} \right\}.$$



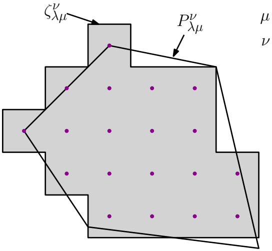
If $P^{\nu}_{\lambda\mu}$ is the hive polytope corresponding to (λ, μ, ν), the number of lattice points in $P^{\nu}_{\lambda\mu} \cap \mathbb{Z}^{\binom{n}{2}}$ is equal to the volume of

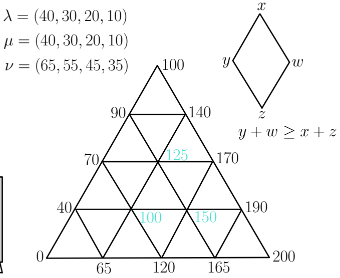
$$\zeta^{\nu}_{\lambda\mu} := \left\{ x \Big| \inf_{y \in P^{\nu}_{\lambda\mu} \cap \mathbb{Z}^{\binom{n}{2}}} \|x - y\|_{\infty} < \frac{1}{2} \right\}.$$



Algorithm

1. Compute the volume V of the the polytope $Q_{\lambda\mu}^{\nu}$ obtained by adding a slack of 2 to each constraint of $P_{\lambda\mu}^{\nu}$.
2. Produce $N = \frac{e^{O(1/\alpha)} \log \frac{1}{\delta}}{\epsilon^2}$ random points from $Q_{\lambda\mu}^{\nu}$, take the nearest lattice point to each, and compute the proportion f of the resulting points that lie in $P_{\lambda\mu}^{\nu}$.
3. Output fV (an estimate of $vol\, \zeta_{\lambda\mu}^{\nu}$).





Proof of correctness

$\alpha = \Omega(1/\log n)$,
$\Delta = \alpha(n^3, n^3 - n^2, \ldots, n^2)$,

$\Delta' = \alpha(\frac{3n^3}{2} + \frac{n^2}{2}, \frac{3n^3}{2} - \frac{n^2}{2}, \ldots, \frac{n^3}{2} + \frac{3n^2}{2})$.

1. $vol\ Q_{\lambda\mu}^{\nu} = e^{O(1/\alpha)} vol \zeta_{\lambda\mu}^{\nu}$ if $(\lambda, \mu, \nu) \in (\Delta, \Delta, \Delta') + LRcone$.

Proof of correctness

$\alpha = \Omega(1/\log n)$,
$\Delta = \alpha(n^3, n^3 - n^2, \ldots, n^2)$,

$\Delta' = \alpha(\frac{3n^3}{2} + \frac{n^2}{2}, \frac{3n^3}{2} - \frac{n^2}{2}, \ldots, \frac{n^3}{2} + \frac{3n^2}{2})$.

1. $vol\ Q_{\lambda\mu}^{\nu} = e^{O(1/\alpha)} vol \zeta_{\lambda\mu}^{\nu}$ if $(\lambda, \mu, \nu) \in (\Delta, \Delta, \Delta') + LRcone$.
2. For a $1 - O(\gamma)$ fraction of all integer points in $LRcone \cap \left\{ \|(\lambda, \mu, \nu)\|_1 \leq \frac{\alpha n^5}{\gamma} \right\}$, $(\lambda, \mu, \nu) \in (\Delta, \Delta, \Delta') +$ LR cone holds.

Proof of correctness

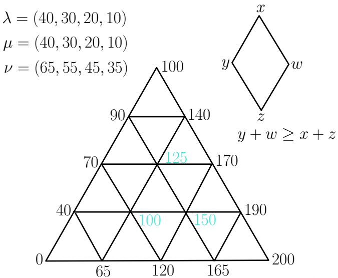
$\alpha = \Omega(1/\log n)$,
$\Delta = \alpha(n^3, n^3 - n^2, \ldots, n^2)$,

$\Delta' = \alpha(\frac{3n^3}{2} + \frac{n^2}{2}, \frac{3n^3}{2} - \frac{n^2}{2}, \ldots, \frac{n^3}{2} + \frac{3n^2}{2})$.

1. $vol\ Q^{\nu}_{\lambda\mu} = e^{O(1/\alpha)} vol \zeta^{\nu}_{\lambda\mu}$ if $(\lambda, \mu, \nu) \in (\Delta, \Delta, \Delta') + LRcone$.
2. For a $1 - O(\gamma)$ fraction of all integer points in $LRcone \cap \left\{\|(\lambda, \mu, \nu)\|_1 \leq \frac{\alpha n^5}{\gamma}\right\}$, $(\lambda, \mu, \nu) \in (\Delta, \Delta, \Delta') +$ LR cone holds.
3. In randomized strongly polynomial time, a nearly random point can be produced from $Q^{\nu}_{\lambda\mu}$, and the volume of $Q^{\nu}_{\lambda\mu}$ can be approximated.

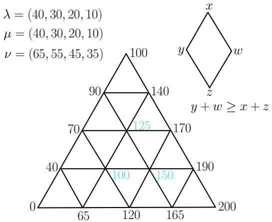
Justifying step 1

The sum of two hives is a hive. Corresponding to $(\Delta, \Delta, \Delta')$, there is a hive with a slack $\frac{\alpha n^2}{2}$ on each constraint. Therefore, corresponding to (λ, μ, ν) in $LRcone + (\Delta, \Delta, \Delta')$, there is a hive with a slack $\frac{\alpha n^2}{2}$ on each constraint, implying that the hive polytope contains a ball of radius $\frac{\alpha n^2}{2}$. So by (a slight modification of) [Kannan-Vempala'97], the number of vertices in $P^{\nu}_{\lambda\mu}$ is at least $e^{-O(1/\alpha)} vol Q^{\nu}_{\lambda\mu}$.



Justifying step 2

If $\|\lambda + \delta_1\|_1 + \|\lambda + \delta_2\|_1 = \|\lambda + \delta_3\|_1$ and $4\|(\delta_1, \delta_2, \delta_3)\|_\infty < \delta\alpha$, then $(\lambda + \delta_1, \mu + \delta_2, \nu + \delta_3) \in LRcone + (1 - \delta)(\Delta, \Delta, \Delta')$ because the minimum norm Lipschitz extension (also called $\infty-$harmonic function) corresponding to $(\delta_1, \delta_2, \delta_3)$ for the equilateral triangle, violates no hive constraint by more than $4\|(\delta_1, \delta_2, \delta_3)\|_\infty$. So on perturbing (λ, μ, ν) to $(\lambda + \delta_1, \mu + \delta_2, \nu + \delta_3)$, the radius of the largest ball in the corresponding hive polytope changes by at most δ.



Theorem

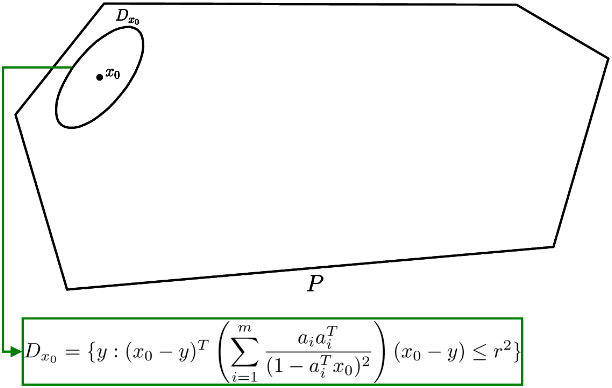
Suppose b is an arbitrary rational vector, A is $m \times d$ *having bitlength poly*(m) *and* $P := \{Ax \leq b\}$*.*

Then vol (P) *can be computed in strongly polynomial time to within* $1 \pm \epsilon$ *with probability* $> 1 - \delta$ *using poly*$(\frac{m}{\epsilon})\log(1/\delta)$ *random bits and poly*$(\frac{m}{\epsilon})\log(1/\delta)$ *arithmetic operations on numbers with bitlength* $\sum_i |\log b_i| + polylog(\frac{d}{\epsilon\delta})$*.*

Proof:

For every constraint hyperplane, find a farthest vertex in strongly polynomial time.

Compute the center of mass x_0 of these points.



Lemma

"Dikin ellipsoid" D_0 at x_0 satisfies $D_0 \subseteq P \subseteq m^{\frac{3}{2}} D_0$.

Perform a linear transformation T to make D_0 the unit ball (in strongly polytime).

$T(P)$ contains the unit ball centered at $T(x_0)$ and is contained in the ball of radius $m^{\frac{3}{2}}$ centered at $T(x_0)$.

Compute $vol T(P)/det T$ using any known $O^*(md^5)$ arithmetic steps and $O^*(d^4)$ bits using Lovász-Vempala'03. □

Log concavity

Conjecture (Okounkov'00)

Viewed as a function from LRcone $\cap\, \mathbb{Z}^{3n}$ *to* $\mathbb{Z}$, $c_{\lambda\mu}^{\nu}$ *is log-concave.*

Theorem (Chindris-Derkson-Weyman'07)

This is false in general!

Approximate log concavity

$\Delta_\epsilon = (n^3, n^3 - n^2, \ldots, n^2)/\epsilon,$

$\Delta'_\epsilon = (\frac{3n^3}{2} + \frac{n^2}{2}, \frac{3n^3}{2} - \frac{n^2}{2}, \ldots, \frac{n^3}{2} + \frac{3n^2}{2})/\epsilon.$

Theorem

If $(\lambda_i, \mu_i, \nu_i) \in LRcone + (\Delta_\epsilon, \Delta_\epsilon, \Delta'_\epsilon)$*, and* $(\lambda_2, \mu_2, \nu_2) = \beta(\lambda_1, \mu_1, \nu_1) + (1-\beta)(\lambda_3, \mu_3, \nu_3)$*, then* $\ln c^{\nu_2}_{\lambda_2\mu_2} \geq \beta \ln c^{\nu_1}_{\lambda_1\mu_1} + (1-\beta) \ln c^{\nu_1}_{\lambda_1\mu_1} - 2\epsilon.$

Proof.

Use Brunn-Minkowski inequality for the volumes of Hive polytopes together with an argument relating the volumes to the number of integer points in them. □

GCT V

The failure of saturation has to be dealt with in testing the positivity of the Kronecker and Plethysm coefficients central to GCT. This issue also arises in testing the positivity of a generalized LR coefficient for an arbitrary semisimple Lie algebra, making the latter a useful prototype.

Theorem (GCT V (Mulmuley-N'07))

Assuming a mathematical conjecture, the positivity of a generalized LR coefficient for an arbitrary semisimple Lie algebra G can be tested in strongly polynomial time.

Classification of Simple Lie algebras

- A_ℓ $SL(\ell+1)$
- B_ℓ $SO(2\ell+1)$
- C_ℓ $Sp(2\ell)$ (symplectic group)
- D_ℓ $SO(2\ell)$
- exceptionals E_6, E_7, E_8, G_2, F_4

Classification of Simple Lie algebras

- A_ℓ $SL(\ell + 1)$
- B_ℓ $SO(2\ell + 1)$
- C_ℓ $Sp(2\ell)$ (symplectic group)
- D_ℓ $SO(2\ell)$
- exceptionals E_6, E_7, E_8, G_2, F_4
- Any semisimple Lie algebra splits into a direct sum of simple Lie algebras.

GCT V

- An LR coefficient for a semisimple algebra is a product of LR coefficients corresponding to its simple factors so the question reduces to testing positivity in the cases of simple Lie algebras.
- For type A, can be tested in polynomial time by linear programming [Knutson-Tao'01], in fact strongly polynomial time [Mulmuley-Sohoni'05].
- Suffices to handle types B, C, D and exceptionals.

BZ polytopes

- For any semisimple Lie algebra and partitions λ, μ, ν, Berenstein and Zelevinsky define a polytope $P^{\nu}_{\lambda\mu}$, the number of integer points in which is $c^{\nu}_{\lambda\mu}$.
- $P^{k\nu}_{k\lambda\, k\mu} = kP^{\nu}_{\lambda\mu}$.
- "Stretching conjecture" of [Loera-McAllister'06] implies: If G is of type B, C or D, and for some $k \in 2\mathbb{N}+1 := \mathbb{M}$, $c^{k\nu}_{k\lambda\, k\mu} > 0$, then $c^{\nu}_{\lambda\mu} > 0$.
- So $P^{\nu}_{\lambda\mu} \cap \mathbb{Z}^{\ell} \neq \phi$ iff $P^{\nu}_{\lambda\mu} \cap \mathbb{Z}^{\ell}_{\mathbb{M}} \neq \phi$.

(As far as we are concerned, $\mathbb{M}$ can be any infinite multiplicatively closed subset of $\mathbb{N}$ that may vary with λ, μ, ν.)

Lemma

- If $\mathbb{Z}^\ell \cap$ *affine span*(P) is nonempty, then
$$P \cap \mathbb{Z}^\ell \neq \phi.$$
- Reason: Let $\mathbb{Z}_{\mathbb{M}}$ denote the subset of rationals whose denominators divide an element of $\mathbb{M}$. If $S :=$ *affine span*(P) is a rational affine subspace, either $S \cap \mathbb{Z}^\ell = \phi$ or $S \cap \mathbb{Z}^\ell_{\mathbb{M}}$ is dense in S.

Algorithm

- Compute the affine span of P using strongly polynomial linear programming.

Algorithm

- Compute the affine span of P using strongly polynomial linear programming.
- Test if $S \cap \mathbb{Z}^\ell$ is empty in strongly polynomial time using the Smith normal form (computed by Kannan-Bachem) as follows.

Algorithm

- Compute the affine span of P using strongly polynomial linear programming.
- Test if $S \cap \mathbb{Z}^{\ell}$ is empty in strongly polynomial time using the Smith normal form (computed by Kannan-Bachem) as follows.
  - Suppose $S := \{Ax = b\}$, where A and b have integer entries.

Algorithm

- Compute the affine span of P using strongly polynomial linear programming.
- Test if $S \cap \mathbb{Z}^\ell$ is empty in strongly polynomial time using the Smith normal form (computed by Kannan-Bachem) as follows.
  - Suppose $S := \{Ax = b\}$, where A and b have integer entries.
  - Compute the Smith normal form $A = UDV$, where U, V are invertible integer matrices and

$$D = diag(d_1, \ldots, d_n),$$

where $d_i \big| d_{i+1}$ are integers.

Algorithm

- Compute the affine span of P using strongly polynomial linear programming.
- Test if $S \cap \mathbb{Z}^\ell$ is empty in strongly polynomial time using the Smith normal form (computed by Kannan-Bachem) as follows.
  - Suppose $S := \{Ax = b\}$, where A and b have integer entries.
  - Compute the Smith normal form $A = UDV$, where U, V are invertible integer matrices and

    $$D = diag(d_1, \dots, d_n),$$

    where $d_i \big| d_{i+1}$ are integers.
  - Solution exists iff $\exists y \in \mathbb{Z}^\ell$ such that $Dy = U^{-1}b$, but D is diagonal so determining this is easy.

Exceptional Lie algebras

For exceptional simple algebras, the cone of 3−tuples (λ, μ, ν) splits into a finite number of subcones, such that within each subcone, $c^{\nu}_{\lambda\mu}$ is a fixed quasi-polynomial function of (λ, μ, ν). Therefore, the LR coefficient can be computed exactly using $O(1)$ arithmetic operations.

Thank You!