Invariant measures and random lattices

Peter Forrester,

M&S, University of Melbourne

Outline

- Invariant measures of Hurwitz and Siegel
- The fundamental domain random lattices, and lattice reduction

Hurwitz and invariant measure

Problems in invariant theory, in particular Cayley's finiteness problem, led Hurwitz to the notion of the invariant measure (now also referred to as the Haar measure) on the continuous matrix groups $SO(N)$ and $U(N)$.

The defining property is

$$\mu(G) = \mu(G_0 G) = \mu(G G_0)$$

With the notation (dA) denoting the product of all independent real and imaginary parts of the differentials of the matrix A, Hurwitz observed that $(R^T dR)$ and $(U^\dagger dU)$ define the invariant measures. Thus, for example, with $R \mapsto R_0 R$, $dR \mapsto R_0 dR$, and so $((R_0 R)^T d(R_0 R)) = (R^T dR)$.

Parametrisation

Hurwitz used Euler angles. In the simplest cases, for $N = 2$, $R \in SO(2)$,

$$R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

$$R^T dR = \begin{bmatrix} 0 & -d\theta \\ d\theta & 0 \end{bmatrix}, \qquad (R^T dR) = d\theta$$

while for $U \in U(2)/(U(1))^2$,

$$U = \begin{bmatrix} \cos\phi & \sin\phi \\ -e^{-i\alpha}\sin\phi & e^{-i\alpha}\cos\phi \end{bmatrix},$$

and

$$U^\dagger dU = -i \begin{bmatrix} \sin^2\phi\, d\alpha & i\gamma \\ -i\bar{\gamma} & \cos^2\phi\, d\alpha \end{bmatrix}$$

where $\gamma = d\phi + i \sin\phi\cos\phi\, d\alpha$ and thus $(U^\dagger dU) = \sin\phi\cos\phi\, d\phi\, d\alpha$.

Hurwitz evaluated $\int (R^T dR)$ and $\int (U^\dagger dU)$.

Parametrisation for SO(3)

The dimension of SO(N) is $N(N-1)/2$. Hence 3 parameters are needed for SO(3). In terms of Euler angles

$$\begin{bmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

where $0 \leq \theta \leq \pi$, $0 \leq \phi, \psi < 2\pi$.

Remark Multiplying shows the final column equals $\begin{bmatrix} \sin\theta\sin\phi \\ \sin\theta\cos\phi \\ \cos\theta \end{bmatrix}$.

This is the spherical coordinate of a unit vector in $\mathbb{R}^3$ with θ the polar angle, ϕ the azimuthal angle.

A calculation gives

$$(R^T dR) = \sin\theta \, d\theta d\phi d\psi.$$

Decomposing the invariant measure (Weyl)

Given $U \in U(N)$, from linear algebra, for $V \in U(N)/(U(1))^N$, we have the eigenvalue/ eigenvector decomposition

$$U = VDV^\dagger,$$

where $D = \mathrm{diag}\,(e^{i\theta_1}, \ldots, e^{i\theta_N})$.

A method based on line elements: $(ds)^2 = \mathrm{Tr}\, dU dU^\dagger$ gives the factorisation

$$(U^\dagger dU) = \prod_{1 \le j < k \le N} |e^{i\theta_k} - e^{i\theta_j}|^2 \, d\theta_1 \cdots d\theta_N \, (V^\dagger dV)$$

Thus, noting $dV\, V^\dagger = -V dV^\dagger$ we have

$$V^\dagger dU V = V^\dagger dV\, D + dD - DV dV^\dagger$$

From this

$$\mathrm{Tr}\, dU dU^\dagger = \sum_{j=1}^{N} (d\theta_j)^2 + 2 \sum_{j<k} |e^{i\theta_k} - e^{i\theta_j}|^2 (((\vec{v}_j^\dagger d\vec{v}_k)^{\mathrm{r}})^2 + ((\vec{v}_j^\dagger d\vec{v}_k)^{\mathrm{i}})^2)$$

Invariant measure for $GL_N(\mathbb{R})$ and $SL_N(\mathbb{R})$

Work of Siegel on the geometry of numbers lead him to consider the invariant measure on $GL_N(\mathbb{R})$,

$$d\mu(M) = \frac{(dM)}{|\det M|^N}$$

Here $(dM) = \prod_{i,j=1}^{N} dM_{i,j}$.

For matrices $A \in SL_N(\mathbb{R})$, Siegel defines the cone λA, $0 < \lambda < 1$, $\lambda A \in GL_N(\mathbb{R})$. From above, the latter has invariant measure equal to the Lebesgue measure (dA). Equivalently, the invariant measure for matrices in $SL_N(\mathbb{R})$ is equal to

$$\delta\Big(1 - \det M\Big)(dM)$$

for $M \in GL_N(\mathbb{R})$.

Singular value decomposition

Jack and Macbeath (1959) considered the singular value decomposition

$$M = O_1 D O_2^T$$

with $O_1, O_2 \in O(N)$ and $D = \mathrm{diag}(\sigma_1, \ldots, \sigma_N)$ of the invariant measure.
Using the method of line elements they obtained the formula

$$(dM) = 2^{-N}(O_1^T dO_1)(O_2^T dO_2)\delta\Big(1-\prod_{l=1}^{N}\sigma_l\Big)\prod_{1\le j<k\le N}(\sigma_j^2-\sigma_k^2)\,d\sigma_1\cdots d\sigma_N$$

Integrating over $O_1, O_2 \in O(N)$ using Hurwitz's result leaves the task of computing the corresponding volume as

$$\int_{R>\sigma_1>\cdots>\sigma_N>0}\delta\Big(1-\prod_{l=1}^{N}\sigma_l\Big)\prod_{1\le j<k\le N}(\sigma_j^2-\sigma_k^2)\,d\sigma_1\cdots d\sigma_N$$

The cutoff R is needed since the measure is not normalisable for large singular values.

Integration method and application

Integration techniques made popular in studies of random matrix products suggest using the Mellin transform. Thus insert t in $\delta\left(t - \prod_{l=1}^{N} \sigma_l\right)$, multiply by t^s, and integrate over t to get

$$\frac{2^{-N}}{N!} \int_0^{R^2} dx_1 \cdots \int_0^{R^2} dx_N \prod_{l=1}^{N} x_l^{s/2-1} \prod_{1 \le j < k \le N} |x_k - x_j|$$

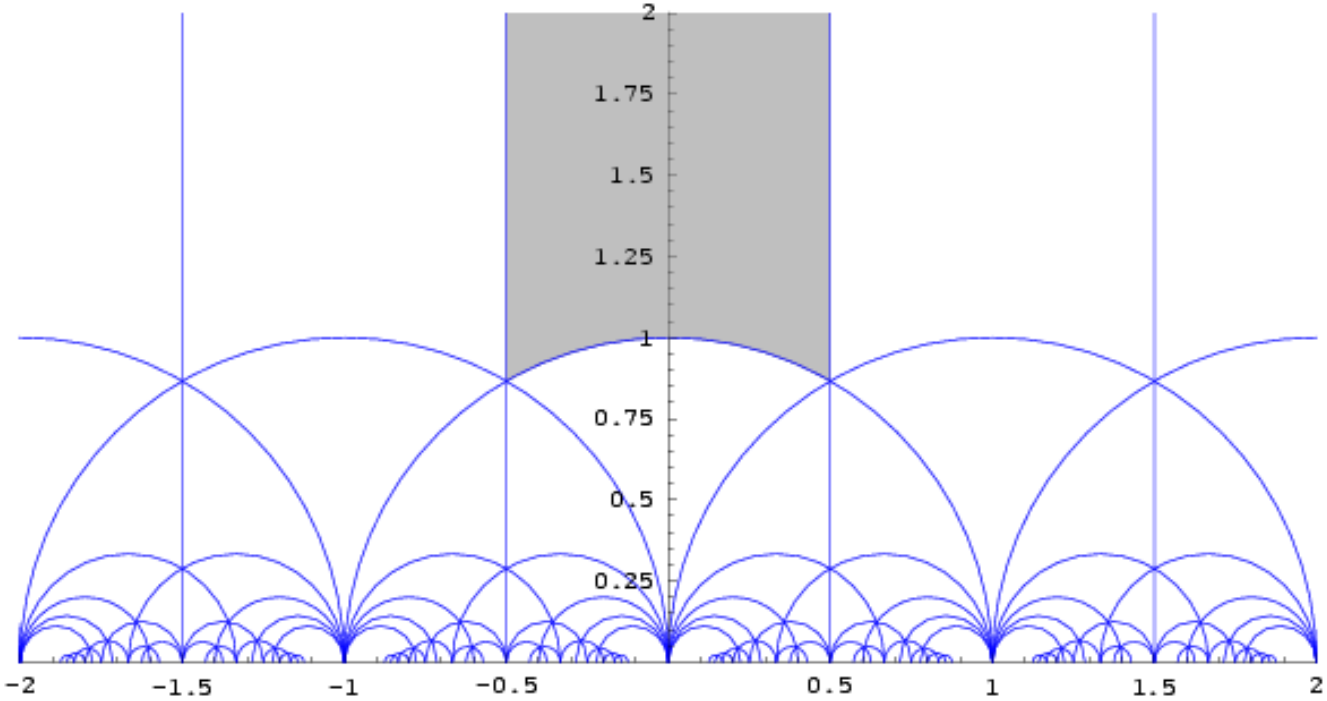
The Selberg integral gives this as a product of gamma functions. Taking the inverse Mellin transform and setting $t = 1$ implies an evaluation proportional to

$$\frac{R^{N^2-N}}{2\pi i} \int_{c-i\infty}^{c+i\infty} R^{Ns} \prod_{j=0}^{N-1} \frac{\Gamma((s+j)/2)}{\Gamma((s+N+1+j)/2)}\, ds$$

$$\propto \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \left(\frac{1}{w}\right)^{[(N+1)/2]} \frac{R^{Nw}}{\prod_{r=1}^{N-1} (w^2 - (N-r)^2)^{[(r+1)/2]}}\, dw.$$

Application: counting matrices in $\mathrm{SL}_N(\mathbb{Z})$

Associated with the quotient $\mathrm{SL}_N(\mathbb{R})/\mathrm{SL}_N(\mathbb{Z})$ is a fundamental domain Γ and a tessellation of $\mathrm{SL}_N(\mathbb{R})$. E.g. $N = 2$



According to Duke, Rudnick and Sarnak (1993)

$$\#\{\gamma : \gamma \in \mathrm{SL}_N(\mathbb{Z}), ||\gamma|| \leq R\} \underset{R\to\infty}{\sim} \frac{1}{\mathrm{vol}\,\Gamma} \int_{||G||\leq R} d\mu(G).$$

The result of Jack and Macbeath tell us the RHS for $||G|| = \sigma_1$ (the operator norm).

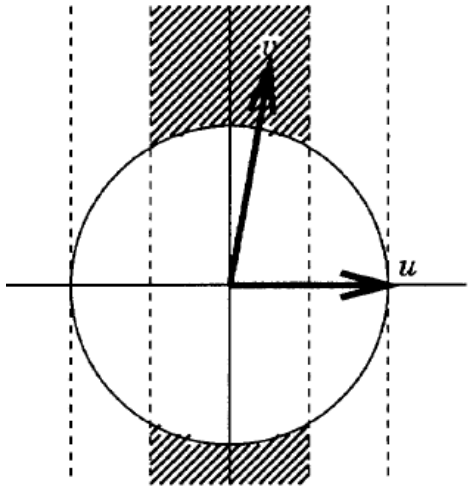
Shortest lattice vector

Basis vectors $\vec{m}_1, \ldots, \vec{m}_n$. Want to choose $(n_1, \ldots, n_N) \neq \vec{0}$ and $\in \mathbb{Z}^N$ such that $\left| \sum_{j=1}^{N} n_j \vec{m}_j \right|$ is minimum.

Question: What is the distribution of the shortest lattice vector when the basis vectors are chosen with invariant measure?

Can answer this question for $N = 2$.

For $N = 2$ it is easy to show that the shortest vector $\mathbf{u}$ and the second shortest, linearly independent vector $\mathbf{v}$ are characterised by the inequalities $||\mathbf{v}|| \geq ||\mathbf{u}||$, $2|\mathbf{u} \cdot \mathbf{v}| \leq ||\mathbf{u}||^2$, the second being equivalent to $||\mathbf{v} + n\mathbf{u}|| \geq ||\mathbf{v}||$ for all $n \in \mathbb{Z}$.



QR decomposition

To align the shortest vector along the x-axis we use the QR decomposition: for $N = 2$

$$\begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} \\ 0 & r_{22} \end{bmatrix}$$

with $r_{11} > 0$ and $r_{22} = 1/r_{11}$. Hence $\mathbf{u} = (r_{11}, 0)$ and $\mathbf{v} = (r_{12}, r_{22})$.

Invariant measure factorises according to

$$d\mu(M) = \delta(1 - \prod_{l=1}^{N} r_{ll}) \prod_{l=1}^{N} r_{ll}^{N-l} (dR)(Q^T dQ).$$

For $N = 2$, integrate over r_{22}, and $(Q^T dQ)$. Leaves $2\pi dr_{11} d_{12}$ — flat measure.

Inequalities for a reduced lattice read $r_{12}^2 + r_{22}^2 \geq r_{11}^2$, $2|r_{12}| \leq r_{11}$.

The distributions

Set $r_{11} = s$, integrate over r_{12}. Shows that the PDF of the length of the shortest lattice vector $\mathbf{u}$ is given by

$$\frac{12}{\pi}\left(\frac{s}{2} - \chi_{s>1}(s^2 - 1/s^2)^{1/2}\right), \qquad 0 < s < (4/3)^{1/4}.$$

Set $(r_{12}^2 + 1/r_{11}^2)^{1/2} = s$. Find the PDF of the second shortest basis vector $\mathbf{v}$ is given by

$$\frac{12}{\pi s}\left((s^4-1)^{1/2}\chi_{1<s<(4/3)^{1/4}} + (2s^2(s^2-(s^4-1)^{1/2})-1)^{1/2})\chi_{(4/3)^{1/4}<s<\infty}\right).$$

Can be illustrated by the following numerical procedure:

1. Generate random matrices M from $\mathrm{SL}_2(\mathbb{R})$ with invariant measure, constrained so that $||M||_{\mathrm{Op}} \leq R$ for some (large) R. For this use the singular value decomposition and Monte Carlo to sample the singular values.
2. Apply Lagrange–Gauss lattice reduction to the columns of M, giving the reduced basis.

Lattice reduction in one-dimension

Euclid GCD

$$a = 210, \quad b = 45$$

$$r_1 = 210 - \left\lfloor \frac{210}{45} \right\rfloor 45 = 210 - 4 \times 45 = 30$$
$$r_2 = 45 - \left\lfloor \frac{45}{30} \right\rfloor 30 = 45 - 1 \times 45 = 15$$
$$r_3 = 30 - \left\lfloor \frac{30}{15} \right\rfloor 15 = 0$$

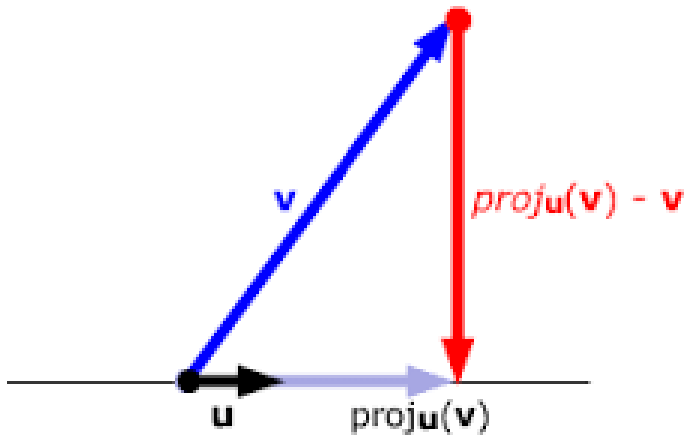
Conclusion: GCD $(210, 45) = 15$

Bézout identity

$$15 = 45 - 1 \times 30 = 45 - 1 \times (210 - 4 \times 45) = -1 \times 210 + 5 \times 45$$

Lattice reduction in two-dimensions: Lagrange-Gauss algorithm

Start with $||\mathbf{u}|| < ||\mathbf{v}||$. Apply orthogonal projection:



Alter $\text{proj}_{\mathbf{u}}(\mathbf{v})$ to equal $\left\lfloor \frac{\mathbf{v}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}} \right\rceil \mathbf{u}$, and compute

$$\mathbf{q} = \mathbf{v} - \left\lfloor \frac{\mathbf{v}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}} \right\rceil \mathbf{u}$$

If $||\mathbf{u}|| < ||\mathbf{q}||$ stop and output reduced basis $\{\mathbf{u}, \mathbf{q}\}$. Else, repeat with input $\mathbf{u}$ and $\mathbf{q}$.

Siegel's mean value theorem

Siegel used the concept of a fundamental domain to prove a mean value theorem:

$$\left\langle \sum_{\substack{\text{lattice points} \\ \vec{y}\setminus\vec{0}}} f(y_1,\ldots,y_N) \right\rangle_{\substack{\text{random} \\ \text{lattice}}} = \int_{\mathbb{R}^N} f(x_1,\ldots,x_N)\, d\vec{x}$$

Consequence Taking $f = \chi_{R-|\vec{y}|}$ implies

$$\left\langle \#\,\text{lattice points in } B_R \right\rangle = \text{vol}\,(B_R)$$

From this, with R_0 such that $\text{vol}\,(B_{R_0}) = 2$, there exists a lattice with shortest vector of length $R_0 \approx O(N^{1/2})$ for large N.

This is the Minkowski lower bound on the biggest shortest vector. Hence realised by a random lattice.

Shortest lattice vector $\mathbf{b}_1$ PDF and Siegel's mean value theorem

R_2: length of 2nd linearly independent vector.
The punctured ball of radius $0 < R < R_2$ will then contain $\pm j\mathbf{b}_1$, where $1 \leq j \leq \left\lfloor \frac{R_2}{||\mathbf{b}_1||} \right\rfloor$.
For small s trial Cs^{d-1} for the PDF of $\mathbf{b}_1$. Let $\Omega(R)$ equal the expected number of lattice vectors in the ball. Then

$$\Omega(R) = 2C \int_0^R \left\lfloor \frac{R}{s} \right\rfloor s^{d-1}\, ds = 2CR^d \int_0^1 \left\lfloor \frac{1}{s} \right\rfloor s^{d-1}\, ds = \frac{2CR^d\zeta(d)}{d}.$$

Siegel's mean value theorem gives $\Omega(R) = \mathrm{Vol}\,(B_R)$. Hence $C = \frac{d}{2\zeta(d)}\mathrm{Vol}\,(B_R)\Big|_{R=1}$. E.g. $d = 3$: