The Longest Increasing Subsequence

Mihir Gupta

Permutation

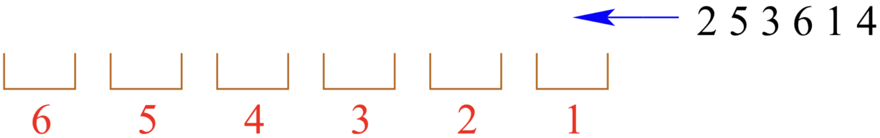
1 3 4 5 2 8 7 6

Permutation

1 3 4 5 2 8 7 6

Length of longest increasing subsequence(LIS) $= 5$

Airplane Boarding



Using a naive model for wait time during airplane boarding the total wait time is equal to the length of the Longest Increasing Subsequence. In the above case it is 3 units.

Standard Young Tableau

Partition of σ where the number in rows and columns are increasing
Example:

1	3	5	8
2	6	9	
4	7		

RSK: Step 1

Permutation: $\{1, 3, 5, 4, 2\}$

Time Step	1
Insertion	1
P-tableau	$\boxed{1}$
Q-tableau	$\boxed{1}$

RSK: Step 2

Permutation: $\{1, 3, 5, 4, 2\}$

Time Step	1	2
Insertion	1	3
P-tableau	$\boxed{1}$	$\boxed{1}\boxed{3}$
Q-tableau	$\boxed{1}$	$\boxed{1}\boxed{2}$

RSK: Step 3

Permutation: $\{1, 3, 5, 4, 2\}$

Time Step	1	2	3
Insertion	1	3	5
P-tableau	$\boxed{1}$	$\boxed{1 \mid 3}$	$\boxed{1 \mid 3 \mid 5}$
Q-tableau	$\boxed{1}$	$\boxed{1 \mid 2}$	$\boxed{1 \mid 2 \mid 3}$

RSK: Step 4

Permutation: $\{1, 3, 5, 4, 2\}$

Time Step	1	2	3	4
Insertion	1	3	5	4
P-tableau	$\begin{array}{\|c\|}\hline 1 \\ \hline\end{array}$	$\begin{array}{\|c\|c\|}\hline 1 & 3 \\ \hline\end{array}$	$\begin{array}{\|c\|c\|c\|}\hline 1 & 3 & 5 \\ \hline\end{array}$	$\begin{array}{\|c\|c\|c\|}\hline 1 & 3 & 4 \\ \hline 5 \\ \cline{1-1}\end{array}$
Q-tableau	$\begin{array}{\|c\|}\hline 1 \\ \hline\end{array}$	$\begin{array}{\|c\|c\|}\hline 1 & 2 \\ \hline\end{array}$	$\begin{array}{\|c\|c\|c\|}\hline 1 & 2 & 3 \\ \hline\end{array}$	$\begin{array}{\|c\|c\|c\|}\hline 1 & 2 & 3 \\ \hline 4 \\ \cline{1-1}\end{array}$

RSK: Step 5

Permutation: $\{1, 3, 5, 4, 2\}$

Time Step	1	2	3	4	5
Insertion	1	3	5	4	2
P-tableau	$\begin{array}{\|c\|}\hline 1 \\ \hline\end{array}$	$\begin{array}{\|c\|c\|}\hline 1 & 3 \\ \hline\end{array}$	$\begin{array}{\|c\|c\|c\|}\hline 1 & 3 & 5 \\ \hline\end{array}$	$\begin{array}{\|c\|c\|c\|}\hline 1 & 3 & 4 \\ \hline 5 \\ \cline{1-1}\end{array}$	$\begin{array}{\|c\|c\|c\|}\hline 1 & 2 & 4 \\ \hline 3 \\ \cline{1-1} 5 \\ \cline{1-1}\end{array}$
Q-tableau	$\begin{array}{\|c\|}\hline 1 \\ \hline\end{array}$	$\begin{array}{\|c\|c\|}\hline 1 & 2 \\ \hline\end{array}$	$\begin{array}{\|c\|c\|c\|}\hline 1 & 2 & 3 \\ \hline\end{array}$	$\begin{array}{\|c\|c\|c\|}\hline 1 & 2 & 3 \\ \hline 4 \\ \cline{1-1}\end{array}$	$\begin{array}{\|c\|c\|c\|}\hline 1 & 2 & 3 \\ \hline 4 \\ \cline{1-1} 5 \\ \cline{1-1}\end{array}$

RSK: Properties

Permutation: $\{1, 3, 5, 4, 2\}$ P-tableau:

1	2	4
3		
5		

Q-tableau:

1	2	3
4		
5		

- Length of top row = len(LIS)
- Length of leftmost column = len(LDS)
- Provides bijection between all pairs of SYT (P and Q tableau) and all permutations of $\{1, 2, \ldots, n\}$. This essentially means that there is a 1-to-1 mapping between pairs of SYT and permutations

RSK Correspondence

λ is a partition of n. For example: $\lambda = (3,2)$ is a partition of 5.
f^{λ} represents the number of SYT with shape λ. For example: $f^{(3,2)} = 5$

$$\begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline 4 & 5 \\ \cline{1-2}\end{array} \quad \begin{array}{|c|c|c|}\hline 1 & 3 & 5 \\ \hline 2 & 4 \\ \cline{1-2}\end{array} \quad \begin{array}{|c|c|c|}\hline 1 & 2 & 4 \\ \hline 3 & 5 \\ \cline{1-2}\end{array} \quad \begin{array}{|c|c|c|}\hline 1 & 3 & 4 \\ \hline 2 & 5 \\ \cline{1-2}\end{array} \quad \begin{array}{|c|c|c|}\hline 1 & 2 & 5 \\ \hline 3 & 4 \\ \cline{1-2}\end{array}$$

Because of bijection between permutations and pairs of SYTs of the RSK algorithm.

$$\sum_{\lambda \vdash n} (f^{\lambda})^2 = n!$$

Hook Length Formula by Frame, Robinson, and Thrall

For a cell $u \in \lambda$, $h(u)$ is the number of cells directly to the right of u, directly below u, and including the cell u. For $\lambda = (3, 2)$, the hook lengths are:

4	3	1
2	1	

Hook Length Formula

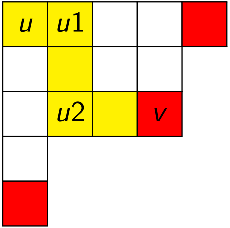
$$f^\lambda = \frac{n!}{\prod_{u \in \lambda} h(u)} = \frac{n!}{H(\lambda)}$$

Example:

$$f^{(3,2)} = \frac{5!}{4 \cdot 3 \cdot 2 \cdot 1 \cdot 1} = 5.$$

Hook Length Formula Proof by Greene et al.

Using probabilistic hook walks and induction.



$$f^{(5,4,4,1,1)} = f^{(4,4,4,1,1)} + f^{(5,4,3,1,1)} + f^{(5,4,4,1)}$$

Sum of probabilities to go to a specific corner v from any u leads us to the desired induction.

$$\sum_u P(u,v) = \frac{H(\lambda)}{H(\lambda - v)} \;\; \& \;\; \sum_v P(u,v) = 1 \implies \frac{n!}{H(\lambda)} = \sum_v \frac{(n-1)!}{H(\lambda - v)}$$

Erdős-Szekeres Theorem

Theorem

For $\sigma \in S_n$ where $n > r \cdot s$, either $len(LIS(\sigma)) > r$ or $len(LDS(\sigma)) > s$

Proof.

$\sigma(1)$	$\sigma(2)$	$\cdots$	$\sigma(r)$
$\sigma(r+1)$	$\sigma(r+2)$	$\cdots$	$\sigma(2r)$
$\vdots$	$\vdots$	$\ddots$	$\vdots$
$\sigma((s-1)r+1)$	$\sigma((s-1)r+2)$	$\cdots$	$\sigma(rs)$

By pigeon-hole principle, either $len(LIS(\sigma)) > r$ or $len(LDS(\sigma)) > s$ ∎

Ulam-Hammersley Problem

The Ulam-Hammersley problem is concered with finding the expected value of the length of the LIS for a permutation σ of length n.

$$l_n = \mathbb{E}_{len(LIS(n))} = \frac{1}{n!} \sum_{\sigma \in S_n} len(LIS(\sigma))$$

Examples:

$$l_1 = 1.00, \quad l_2 = 1.50, \quad \dots \quad l_9 = 4.06, \quad l_{10} = 4.33$$

First Bounds - Hammersley

$$\frac{\pi}{2} \leq \lim_{n\to\infty} \frac{l_n}{\sqrt{n}} \leq e$$

Proof for lower bound uses a Poisson process for projected distances and strong law of large numbers.

Upper bound is established by calculating the expected number of such subsequences, applying Stirling's approximation and Markov's inequality to refine probabilistic bounds, and strategically choosing subsequence lengths.

Logan-Shepp & Vershik-Kerov

Lemma (Logan and Shepp, 1977)

$$\lim_{n\to\infty} \frac{l_n}{\sqrt{n}} \geq 2$$

Lemma (Vershik and Kerov, 1977)

$$\lim_{n\to\infty} \frac{l_n}{\sqrt{n}} \leq 2$$

Theorem

The limit of the expected length of the LIS for random permutations exists and equals 2:

$$\lim_{n\to\infty} \frac{l_n}{\sqrt{n}} = 2$$

Baik-Deift-Johansson Theorem

Theorem (BDJ)

For random (uniform) $\sigma \in S_n$ and all $t \in \mathbb{R}$,

$$\lim_{n\to\infty} Prob\left(\frac{is_n(\sigma) - 2\sqrt{n}}{n^{1/6}} \leq t\right) = F(t),$$

where $F(t)$ is the Tracy-Widom distribution.

Provides the scale of fluctuations which are of order $n^{1/6}$.

Precise Expectation of LIS Length

The expectation $E(n)$ can be expressed more precisely as:

$$E(n) = 2\sqrt{n} + \alpha n^{1/6} + o(n^{1/6}),$$

where $\alpha = \int t dF(t) = -1.7710868074...$