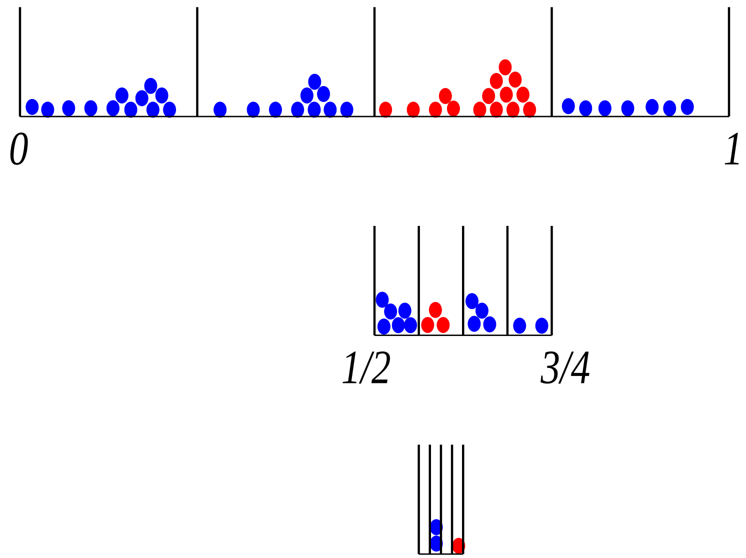
An Introduction to the Contraction Method

Ralph Neininger
Institute for Mathematics
Johann Wolfgang Goethe-Universität
Frankfurt am Main

Bucket Selection

Given: $U_1, \dots, U_n$ indep, unif$[0,1]$ distributed.
Problem: Find element with rank $k \in \{1, \dots, n\}$.



X_n: number of items distributed

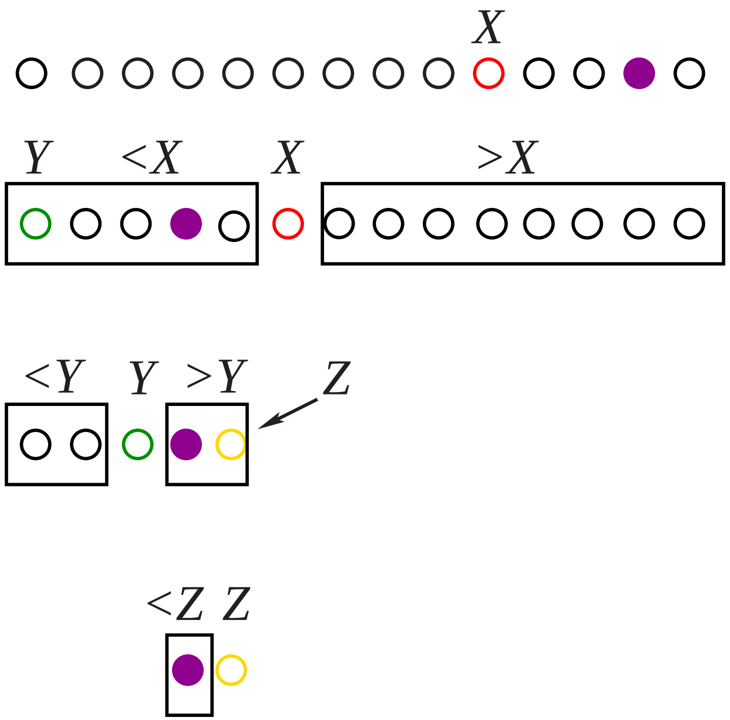
$$X_n \stackrel{d}{=} X_{I_n} + n$$

$I_n = \#$ items in relevant bucket

Quickselect

Given: $x_1, \ldots, x_n \in \mathbb{R}$.

Problem: Find element with rank $k \in \{1, \ldots, n\}$.



X_n: number of key comparisons

$$X_n \stackrel{d}{=} X_{I_n} + n - 1, \quad n \geq 2.$$

I_n = length of relevant sublist.

For $k = 1$: $I_n \stackrel{d}{=} \mathrm{unif}\{0, \ldots, n-1\}$

Quickselect: Analysis for $k = 1$

$$X_n \stackrel{d}{=} X_{I_n} + n - 1, \quad n \geq 2, \quad (X_0 = X_1 = 0).$$

$$I_n \stackrel{d}{=} \mathrm{unif}\{0, \ldots, n-1\}$$

I_n independent of $X_0, \ldots, X_{n-1}$.

Scaling

$$Y_n := \frac{X_n}{n} \stackrel{d}{=} \frac{X_{I_n}}{n} + \frac{n-1}{n} = \frac{I_n}{n} \cdot \frac{X_{I_n}}{I_n} + \frac{n-1}{n} = \underbrace{\frac{I_n}{n}}_{\to U} Y_{I_n} + \underbrace{\frac{n-1}{n}}_{\to 1}.$$

With $n \to \infty$:

$$\frac{X_n}{n} = Y_n \to Y \stackrel{d}{=} UY + 1$$

with U, Y independent and $U \stackrel{d}{=} \mathrm{unif}[0, 1]$.

Quickselect: Analysis for $k = 1$

$$\frac{X_n}{n} = Y_n \to Y \overset{d}{=} UY + 1$$

with U, Y independent and $U \overset{d}{=} \text{unif}[0,1]$.

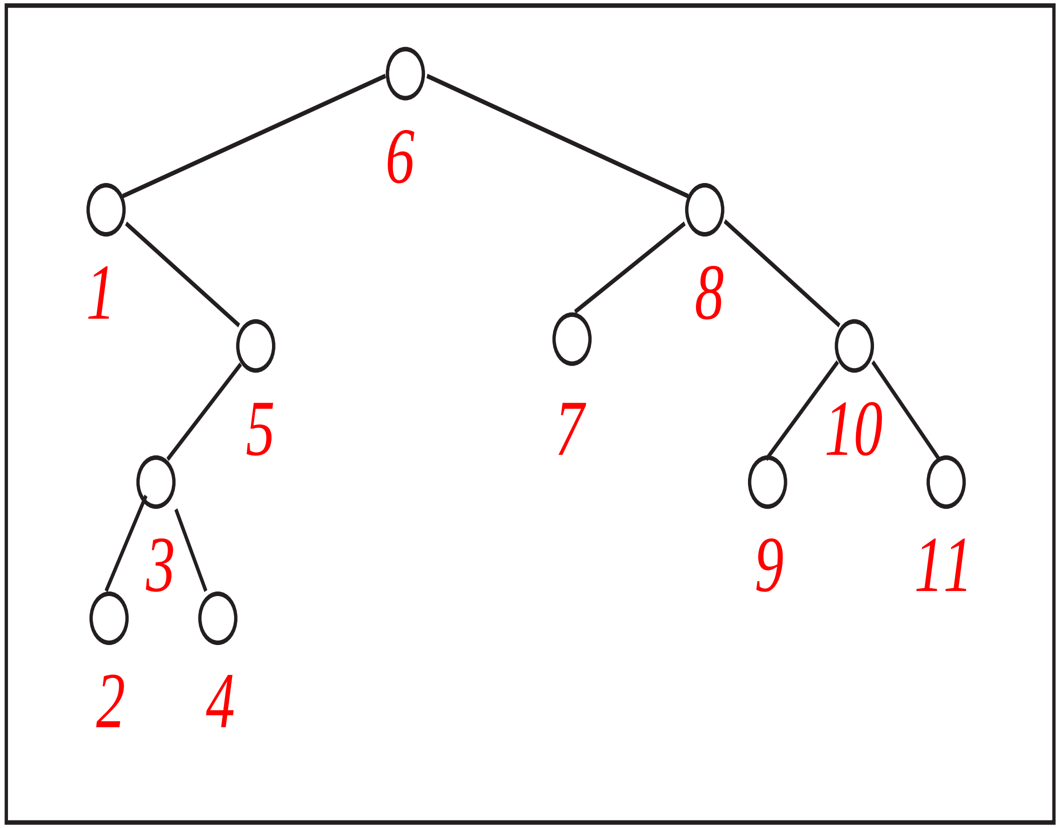
$$\begin{aligned}
\mathbb{E}\, Y &= \mathbb{E}\,[UY] + 1 = \frac{1}{2}\mathbb{E}\, Y + 1 \quad \Rightarrow \quad \mathbb{E}\, Y = 2, \\
\mathbb{E}\left[Y^2\right] &= \mathbb{E}\left[U^2Y^2 + 2UY + 1\right] = \frac{1}{3}\mathbb{E}\left[Y^2\right] + 3 \quad \Rightarrow \quad \mathbb{E}\left[Y^2\right] = \frac{9}{2}, \\
\text{Var}(Y) &= \frac{1}{2}.
\end{aligned}$$

Hence, this suggests

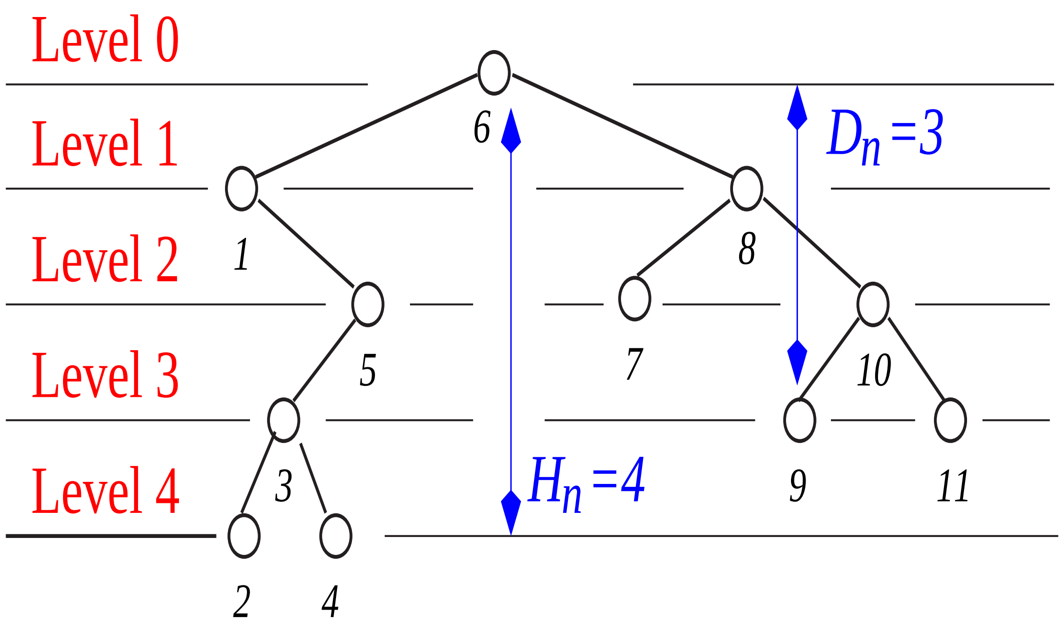
$$\begin{aligned}
\mathbb{E}\, X_n &= \mathbb{E}\,[nY_n] \sim n\mathbb{E}\, Y = 2n, \\
\text{Var}(X_n) &= \text{Var}(nY_n) \sim n^2\text{Var}(Y) = \frac{1}{2}n^2.
\end{aligned}$$

Binary search tree

Given numbers: 6, 1, 8, 7, 5, 3, 10, 2, 11, 4, 9.



Quantities in BST



D_n — depth = distance root to n-th inserted node

$H_n = \max_{1 \le j \le n} D_j$ — height

$Q_n = \sum_{1 \le j \le n} D_j$ — internal path length

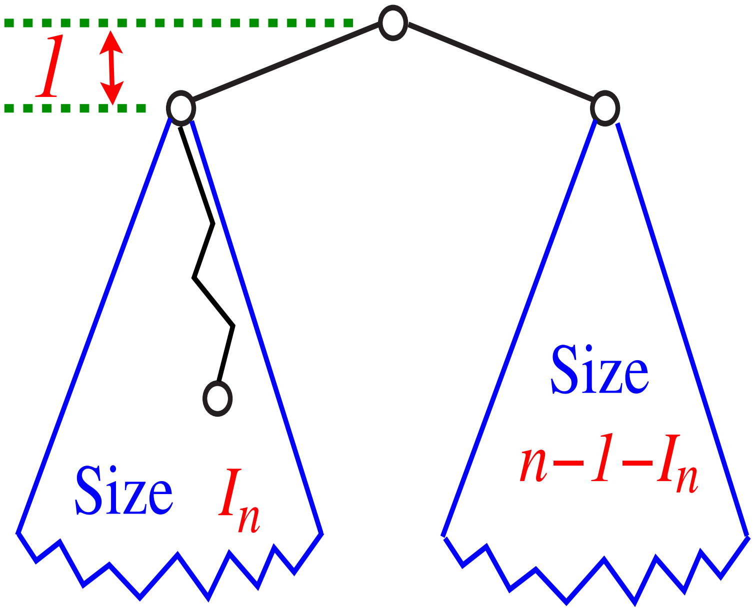
Random binary search tree

Model of randomness:

All permutations of $1, \ldots, n$ equally likely.

Equivalent: Use $U_1, \ldots, U_n$ i.i.d. unif$[0, 1]$.

Internal path length



Internal path length Q_n:

$$Q_n \stackrel{d}{=} Q^*_{I_n} + Q^{**}_{n-1-I_n} + n - 1$$

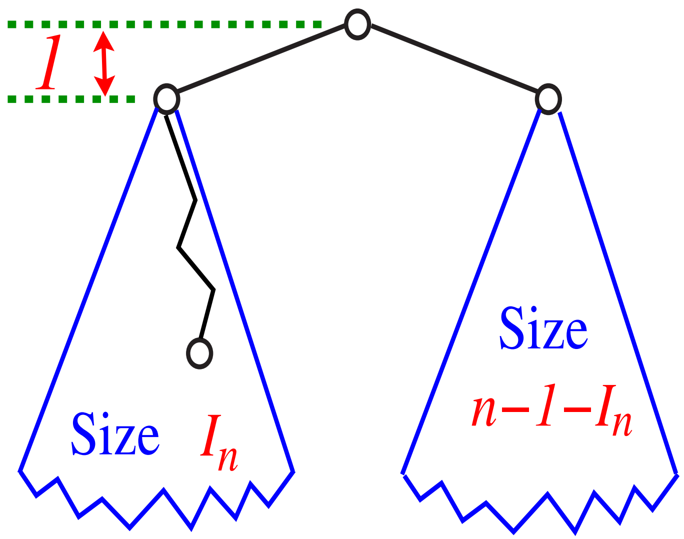
$Q^*_0, \dots, Q^*_{n-1}, Q^{**}_0, \dots, Q^{**}_{n-1}, I_n$ indep.,

$Q^*_j \stackrel{d}{=} Q^{**}_j \stackrel{d}{=} Q_j$,

$I_n \stackrel{d}{=} \text{unif}\{0, \dots, n-1\}$.

This needs a proof!

Internal path length Q_n



$$Q_n \stackrel{d}{=} Q^*_{I_n} + Q^{**}_{n-1-I_n} + n - 1 =: Z_n,$$

$Q^*_0, \dots, Q^*_{n-1}, Q^{**}_0, \dots, Q^{**}_{n-1}, I_n$ indep.,

$$Q^*_j \stackrel{d}{=} Q^{**}_j \stackrel{d}{=} Q_j,$$

$$I_n \stackrel{d}{=} \mathrm{unif}\{0, \dots, n-1\}.$$

Sufficient: $\mathbb{P}(Q_n = j) = \mathbb{P}(Z_n = j)$ for all $j \in \mathbb{N}$.

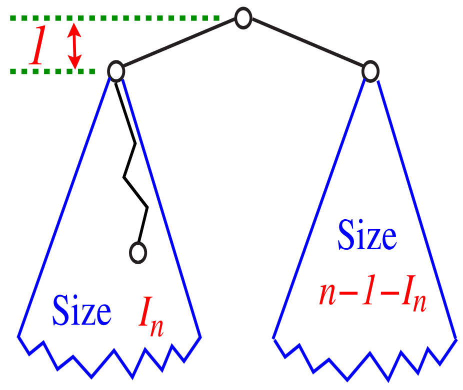
Show: For all $j \in \mathbb{N}$, $k \in \{0, \dots, n-1\}$:

$$\mathbb{P}(Q_n = j \mid I_n = k) = \mathbb{P}(Z_n = j \mid I_n = k).$$

[Total probability theorem yields:

$$\begin{aligned} \mathbb{P}(Q_n = j) &= \sum_k \mathbb{P}(I_n = k)\, \mathbb{P}(Q_n = j \mid I_n = k) \\ &= \sum_k \mathbb{P}(I_n = k)\, \mathbb{P}(Z_n = j \mid I_n = k) = \mathbb{P}(Z_n = j).] \end{aligned}$$

Proof of the recurrence



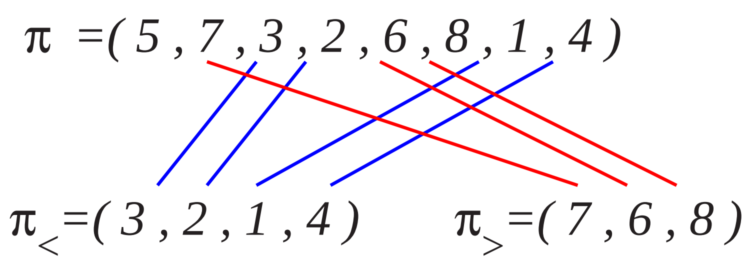
To prove:

$$\mathbb{P}(Q_n = j \,|\, I_n = k)$$

$$= \mathbb{P}(Q_k^* + Q_{n-1-k}^{**} + n - 1 = j).$$

We show: Given $\{I_n = k\}$ we have:

a) $\mathcal{T}_1$ and $\mathcal{T}_2$ independent, b) $\mathcal{T}_1$ and $\mathcal{T}_2$ random BSB.

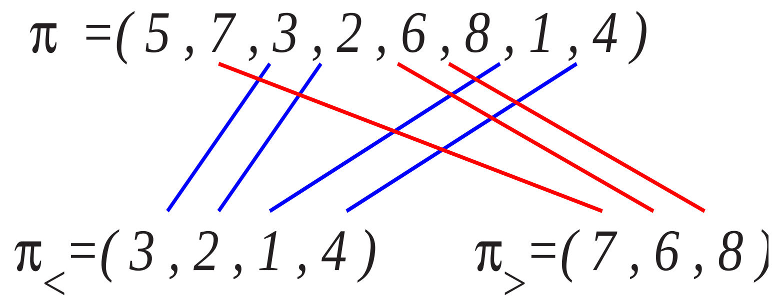


π equiprobable permutation in S_n.

Construct $\pi_<$ and $\pi_>$

Proof of the recurrence II

π equiprobable in S_n.



We show: Given $\pi_1 = k + 1$:

$\pi_<$ and $\pi_>$ independent and equiprobable on S_k and S_{n-1-k}. resp.

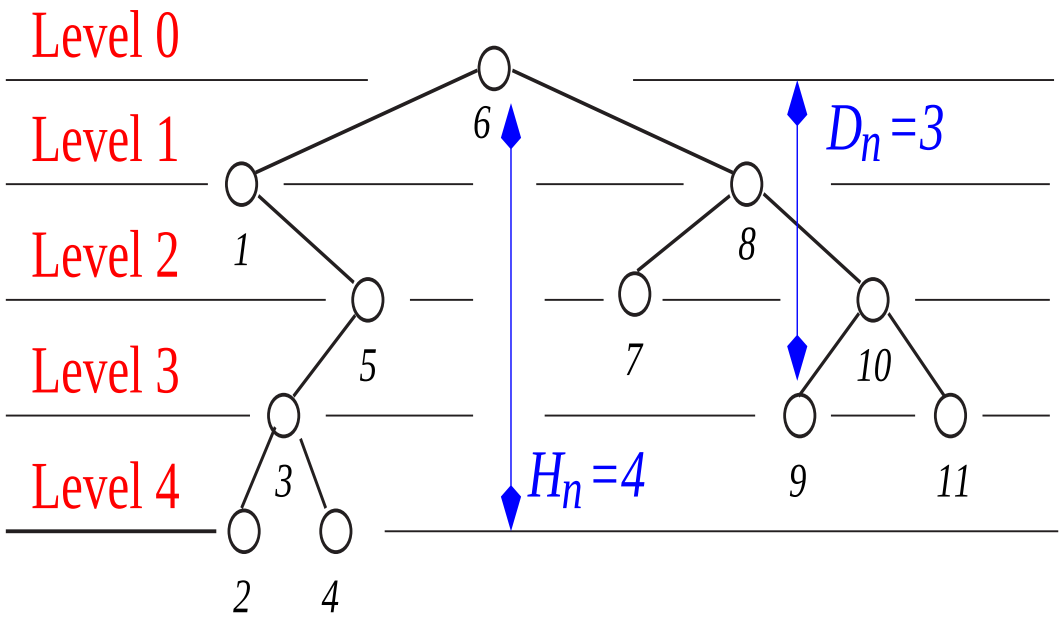
For arbitrary $\sigma \in S_k$:

$$\mathbb{P}(\pi_< = \sigma \,|\, \pi_1 = k+1) = \frac{1}{1/n} \frac{\binom{n-1}{k}(n-1-k)!}{n!} = \frac{1}{k!}$$

Given $\pi_1 = k + 1$ hence $\pi_<$ equiprobable on S_k.

Second assertion similar.

Other BST recurrences



$$Q_n \stackrel{d}{=} Q^{(1)}_{I_n} + Q^{(2)}_{n-1-I_n} + n - 1$$

$$H_n \stackrel{d}{=} H^{(1)}_{I_n} \vee H^{(2)}_{n-1-I_n} + 1$$

$$D_n \stackrel{d}{=} \mathbf{1}_{A_n} D_{I_n} + \mathbf{1}_{A_n^c} D_{n-1-I_n} + 1$$

Expected internal path length

$$Q_n \stackrel{d}{=} Q^*_{I_n} + Q^{**}_{n-1-I_n} + n - 1, \quad I_n \stackrel{d}{=} \text{unif}\{0, \dots, n-1\}.$$

$$\begin{aligned} q_n := \mathbb{E}\, Q_n &= \mathbb{E}\left[Q^*_{I_n} + Q^{**}_{n-1-I_n} + n - 1\right] \\ &= \frac{1}{n}\sum_{k=0}^{n-1} \mathbb{E}\left[Q^*_k + Q^{**}_{n-1-k} + n - 1\right] \\ &= \frac{1}{n}\sum_{k=0}^{n-1} (q_k + q_{n-1-k} + n - 1) \\ &= n - 1 + \frac{2}{n}\sum_{k=0}^{n-1} q_k. \end{aligned}$$

Solves easily:

$$q_n = 2(n+1)\mathcal{H}_n - 4n = 2n\log(n) + (2\gamma - 4)n + o(n).$$

[$\mathcal{H}_n := \sum_{i=1}^{n} 1/i$ harmonic numbers.]

Rescaling

$Q_n \stackrel{d}{=} Q^*_{I_n} + Q^{**}_{n-1-I_n} + n - 1, \quad q_n = 2n\log(n) + cn + o(n).$
Scaling

$$Y_n := \frac{Q_n - q_n}{n}.$$

Then

$$\begin{aligned} Y_n &\stackrel{d}{=} \frac{1}{n}\left(Q^*_{I_n} + Q^{**}_{n-1-I_n} + n - 1 - q_n\right) \\ &= \frac{1}{n}\left(I_n \frac{Q^*_{I_n} \pm q_{I_n}}{I_n} + (n-1-I_n)\frac{Q^{**}_{n-1-I_n} \pm q_{n-1-I_n}}{n-1-I_n} + n - 1 - q_n\right) \\ &= \underbrace{\frac{I_n}{n}}_{\to U} Y^*_{I_n} + \underbrace{\frac{n-1-I_n}{n}}_{\to 1-U} Y^{**}_{n-1-I_n} + b^{(n)} \end{aligned}$$

with

$$b^{(n)} = \frac{1}{n}\left(q_{I_n} + q_{n-1-I_n} - q_n + n - 1\right).$$

Rescaling II

$$q_n = 2n\log(n) + cn + o(n).$$

$$\begin{aligned}
b^{(n)} &= \frac{1}{n}\Big(q_{I_n} + q_{n-1-I_n} - q_n + n - 1\Big). \\
&= \frac{1}{n}\Big(2I_n\log(I_n) + cI_n + 2(n-1-I_n)\log(n-1-I_n) + c(n-1-I_n) \\
&\qquad - 2n\log(n) - cn + o(n) + n - 1\Big) \\
&= \frac{1}{n}\Big(2I_n\log(I_n) \qquad + 2(n-1-I_n)\log(n-1-I_n) \\
&\qquad - 2(I_n + (n-1-I_n) + 1)\log(n) + o(n) + n - 1\Big) \\
&= 2\frac{I_n}{n}\log\left(\frac{I_n}{n}\right) + 2\frac{n-1-I_n}{n}\log\left(\frac{n-1-I_n}{n}\right) + 1 + o(1) \\
&\to 2U\log(U) + 2(1-U)\log(1-U) + 1 =: g(U)
\end{aligned}$$

Rescaling: Summary

$Q_n \stackrel{d}{=} Q^*_{I_n} + Q^{**}_{n-1-I_n} + n - 1, \quad q_n = 2n\log(n) + cn + o(n).$
Scaling

$$Y_n := \frac{Q_n - q_n}{n}.$$

Then

$$Y_n = \underbrace{\frac{I_n}{n}}_{\to U} Y^*_{I_n} + \underbrace{\frac{n-1-I_n}{n}}_{\to 1-U} Y^{**}_{n-1-I_n} + \underbrace{b^{(n)}}_{\to g(U)}$$

with

$$g(U) = 2U\log(U) + 2(1-U)\log(1-U) + 1$$

Hence, this suggests

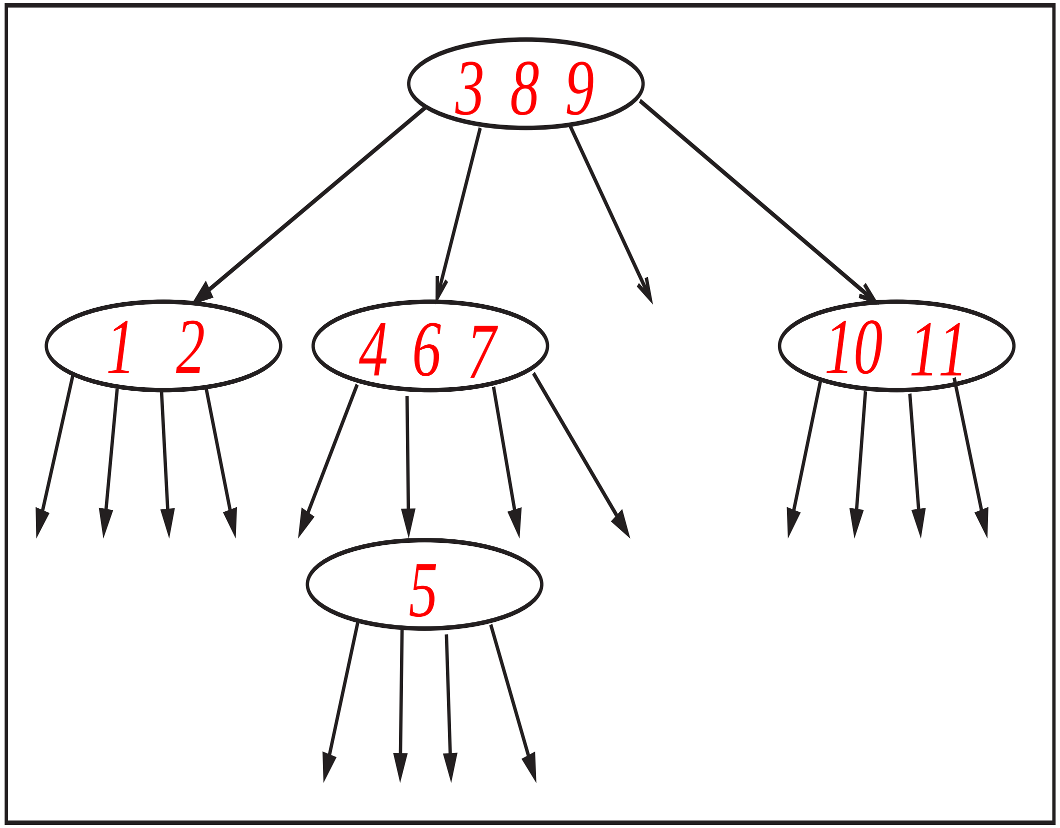
$$Y_n \to Y \stackrel{d}{=} UY^* + (1-U)Y^{**} + g(U),$$

with Y^*, Y^{**}, U independent, $Y \stackrel{d}{=} Y^* \stackrel{d}{=} Y^{**}$.

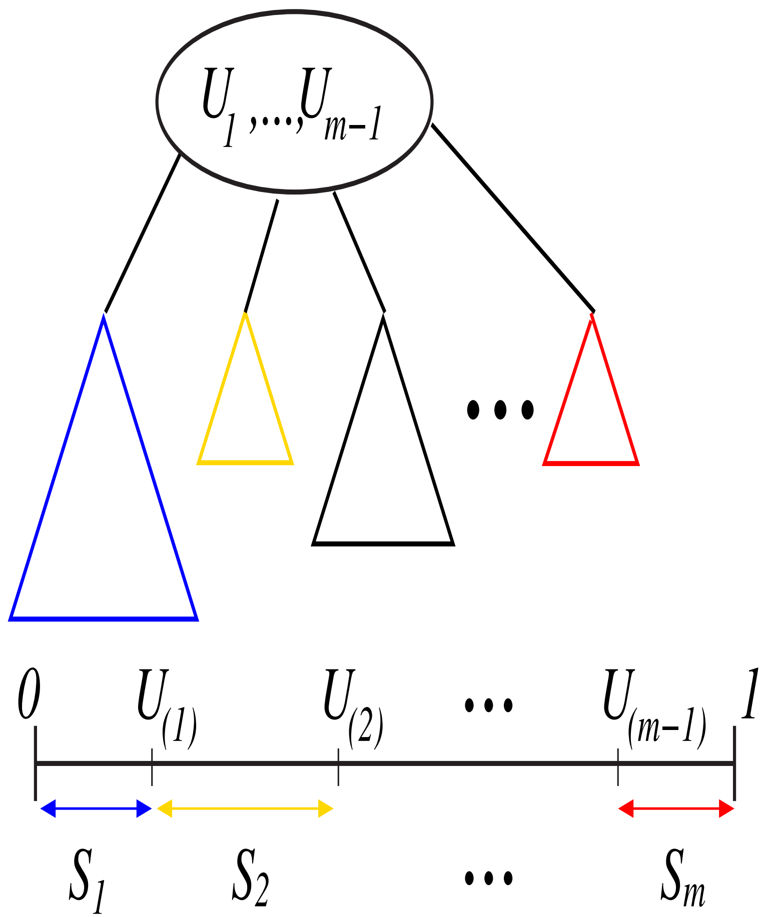
m-ary search trees

Example: $m = 4$

List of data: $8, 3, 9, 6, 2, 1, 11, 7, 10, 4, 5.$



m-ary search tree



Data: $U_1, \dots, U_n$ i.i.d. unif$[0,1]$
(or uniform permutation)

Sizes of subtrees:

$$I^{(n)} \stackrel{d}{=} M(n-m+1, S_1, \dots, S_m),$$

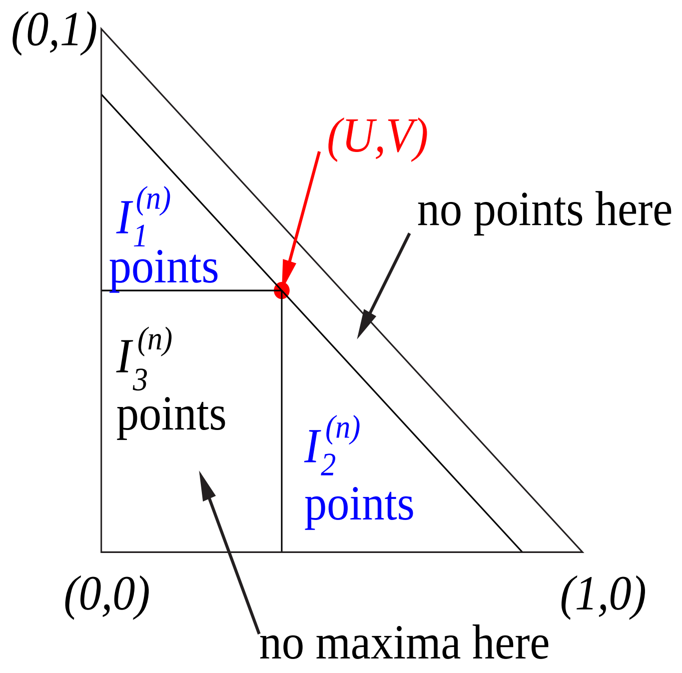
Space needed
(number of int. nodes):

$$X_n \stackrel{d}{=} \sum_{r=1}^{m} X^{(r)}_{I_r^{(n)}} + 1, \quad n \geq m.$$

$$X_0 = 0, X_1 = \cdots = X_{m-1} = 1.$$

Maxima in right triangles

Data: $U_1, \ldots, U_n$ indep. unif. in right triangle



(U, V) has maximal sum of coordinates.

$$X_n \stackrel{d}{=} X^{(1)}_{I_1^{(n)}} + X^{(2)}_{I_2^{(n)}} + 1, \quad n \geq 2.$$

General recursion

$$X_n \stackrel{d}{=} \sum_{r=1}^{K} A_r(n) X^{(r)}_{I^{(n)}_r} + b_n, \quad n > n_0.$$

- $K \geq 1$ Number of subproblems (also $K = K_n$).
- $X^{(r)}_n \stackrel{d}{=} X_n$ (recursive).
- $I^{(n)} = (I^{(n)}_1, \ldots, I^{(n)}_K)$ Sizes of subproblems.
- $\left(X^{(1)}_n\right), \ldots, \left(X^{(K)}_n\right), (A_1(n), \ldots, A_K(n), b_n, I^{(n)})$ independent.

Contraction method

Rösler (1991, 1992)
Rachev and Rüschendorf (1995)

$$X_n \stackrel{d}{=} \sum_{r=1}^{K} A_r(n) X^{(r)}_{I_r^{(n)}} + b_n, \qquad n > n_0.$$

Scaling

$$Y_n := \frac{X_n - \mu(n)}{\sigma(n)}.$$

Then

$$Y_n \stackrel{d}{=} \sum_{r=1}^{K} A_r^{(n)} Y^{(r)}_{I_r^{(n)}} + b^{(n)}, \qquad n > n_0,$$

with

$$A_r^{(n)} = \frac{\sigma(I_r^{(n)})}{\sigma(n)} A_r(n),$$

$$b^{(n)} = \frac{1}{\sigma(n)} \Big(b_n - \mu(n) + \sum_{r=1}^{K} A_r(n) \mu(I_r^{(n)}) \Big).$$

Convergence

Idea:

$$
\begin{array}{ccccccc}
Y_n & \stackrel{d}{=} & \sum_{r=1}^{K} & A_r^{(n)} & Y_{I_r^{(n)}}^{(r)} & + & b^{(n)} \\
\downarrow & & & \downarrow & \downarrow & & \downarrow \\
Y & \stackrel{d}{=} & \sum_{r=1}^{K} & A_r^{*} & Y^{(r)} & + & b^{*}
\end{array}
$$

$$
\left.\begin{array}{l} A_r^{(n)} \longrightarrow A_r^{*} \\ b^{(n)} \longrightarrow b^{*} \end{array}\right\} \Longrightarrow Y_n \longrightarrow Y.
$$

Limit map:

$$
\begin{aligned}
T : \mathcal{M} &\to \mathcal{M} \\
\nu &\mapsto \mathcal{L}\Big(\sum_{r=1}^{K} A_r^{*} Z^{(r)} + b^{*}\Big)
\end{aligned}
$$

with $(A_1^*, \ldots, A_K^*, b^*)$, $Z^{(1)}, \ldots, Z^{(K)}$ independent, $Z^{(r)} \stackrel{d}{=} \nu$.

The minimal ℓ_p metric

$$\begin{aligned}
\mathcal{M} \quad &: \quad \text{Space of probability measures on } \mathbb{R}. \\
\mathcal{M}_p \quad &:= \quad \left\{ \mu \in \mathcal{M} : \|\mu\|_p := \left(\int |x|^p d\mu(x) \right)^{1/p} < \infty \right\}, \qquad p \geq 1, \\
\mathcal{M}_2(0) \quad &:= \quad \left\{ \mu \in \mathcal{M}_2 : \mathbb{E}\,\mu := \int x d\mu(x) = 0 \right\}.
\end{aligned}$$

For r.v. X with distribution $\mathcal{L}(X) = \mu$ we have $\|X\|_p = \|\mu\|_p$.

Definition: The minimal ℓ_p metric ($p \geq 1$ fixed) is given by

$$\begin{aligned}
\ell_p : \mathcal{M}_p \times \mathcal{M}_p \quad &\to \quad [0, \infty) \\
(\mu, \nu) \quad &\mapsto \quad \inf\{\|X - Y\|_p : \mathcal{L}(X) = \mu, \mathcal{L}(Y) = \nu\}
\end{aligned}$$

The minimal ℓ_p metric II

$$\ell_p(\mu, \nu) = \inf\{\|X - Y\|_p : \mathcal{L}(X) = \mu, \mathcal{L}(Y) = \nu\}$$

How can X, Y with distributions μ, ν be constructed so that $\|X - Y\|_p$ is as small as possible?

1st Step: Construct X with $\mathcal{L}(X) = \mu$: Define distribution function of X and its (generalized) inverse by

$$\begin{aligned} F_X(x) &:= \mathbb{P}(X < x) = \mathcal{L}(X)((-\infty, x)), \quad x \in \mathbb{R}, \\ F_X^{-1}(t) &:= \sup\{x \in \mathbb{R} : F(x) \le t\}, \quad t \in \mathbb{R}. \end{aligned}$$

Well-known fact: For a unif$[0,1]$ r.v. U we have

$$\mathcal{L}\left(F_X^{-1}(U)\right) = \mathcal{L}(X).$$

The minimal ℓ_p metric — optimal couplings

2^{nd} step: Use the same unif$[0,1]$ r.v. U for both, i.e.

$$X = F_\mu^{-1}(U), \quad Y = F_\nu^{-1}(U). \tag{1}$$

Definition: A vector (X, Y) with $\mathcal{L}(X) = \mu$, $\mathcal{L}(Y) = \nu$ and

$$\ell_p(\mu, \nu) = \|X - Y\|_p$$

is called an optimal coupling of $\mu, \nu \in \mathcal{M}_p$.

Theorem: Optimal coupling do always exist. For $\mu, \nu \in \mathcal{M}_p$ optimal couplings are given by (1).

Corollary: We have

$$\ell_p(\mu, \nu) = \left(\int_0^1 \left| F_\mu^{-1}(u) - F_\nu^{-1}(u) \right|^p du \right)^{1/p}.$$

The minimal ℓ_p metric — properties

Corollary: $(\mathcal{M}_p, \ell_p)$ is a metric space.

Proof.

- $\ell_p(\mu,\nu) = 0 \Leftrightarrow \mu = \nu$: OK. symmetric: OK.

- Triangle inequality: For given $\mu, \nu, \rho \in \mathcal{M}_p$:

Define

$$X = F_\mu^{-1}(U), \quad Y = F_\nu^{-1}(U), \quad Z = F_\rho^{-1}(U)$$

with a unif$[0,1]$ r.v. U.

(X,Y), (Y,Z), (Y,Z) optimal coupl. of (μ,ν), (ν,ρ), and (ν,ρ) resp.

Hence

$$\ell_p(\mu,\nu) = \|X-Y\|_p \le \|X-Z\|_p + \|Z-Y\|_p = \ell_p(\mu,\nu) + \ell_p(\nu,\rho). \quad \clubsuit$$

The minimal ℓ_p metric — properties

Corollary: $(\mathcal{M}_p, \ell_p)$ is complete.
Proof.
Cauchy sequence (μ_n) in $\mathcal{M}_p$.

Represent μ_n by r.v. X_n such that all pairs are opt. coupl.

$\|X_n - X_m\|_p = \ell_p(\mu_n, \mu_m) \Rightarrow (X_n)$ is Cauchy sequence in $L^p(\Omega, \mathcal{F}, \mathbb{P})$.

$L^p(\Omega, \mathcal{F}, \mathbb{P})$ is complete $\Rightarrow X_n \xrightarrow{L_p} X \in L^p(\Omega, \mathcal{F}, \mathbb{P})$.

$\Rightarrow \mu_n \xrightarrow{\ell_p} \mathcal{L}(X) \in \mathcal{M}_p$. ♣.

The minimal ℓ_p metric — properties

Corollary: For $\mu_n, \mu \in \mathcal{M}_p$ with $\ell_p(\mu_n, \mu) \to 0$:

$$\mu_n \xrightarrow{w} \mu, \quad \|\mu_n\|_p \to \|\mu\|_p \quad (n \to \infty).$$

Proof.
Choose $(X_n), X$ such that all pairs are optimal couplings.

Then $\|X_n - X\|_p \to 0$.

L^p convergence implies convergence in distribution.

Moreover

$$\Big| \|\mu_n\|_p - \|\mu\|_p \Big| = \Big| \|X_n\|_p - \|X\|_p \Big| \le \|X_n - X\|_p \to 0 \quad \clubsuit$$

Lipschitz continuity on $(\mathcal{M}_p, \ell_p)$

Theorem: Assume that $(A_1, \ldots, A_k, b)$ are L^p-integrable r.v.,

$$\begin{aligned} T : \mathcal{M}_p &\to \mathcal{M}_p \\ \mu &\mapsto \mathcal{L}\left(\sum_{r=1}^{K} A_r Z^{(r)} + b\right), \end{aligned}$$

where $(A_1, \ldots, A_k, b), Z^{(1)}, \ldots, Z^{(K)}$ are indep. and $\mathcal{L}(Z^{(r)}) = \mu$.

Then, for all $\mu, \nu \in \mathcal{M}_p$,

$$\ell_p(T(\mu), T(\nu)) \leq \left(\sum_{r=1}^{K} \|A_r\|_p\right) \ell_p(\mu, \nu).$$

If $\sum_{r=1}^{K} \|A_r\|_p < 1$ then T is a contraction on $(\mathcal{M}_p, \ell_p)$.

Proof

$$T(\mu) = \mathcal{L}\left(\sum_{r=1}^{K} A_r Z^{(r)} + b\right), \quad \mu \in \mathcal{M}_p$$

$\|T(\mu)\|_p < \infty$, hence $T(\mathcal{M}_p) \subset \mathcal{M}_p$.

Let $\mu, \nu \in \mathcal{M}_p$.

There exists opt. coupl. $\left(Z^{(1)}, W^{(1)}\right)$ of μ and ν.

Let $\left(Z^{(1)}, W^{(1)}\right), \ldots, \left(Z^{(K)}, W^{(K)}\right)$ be i.i.d.
and independent of $(A_1, \ldots, A_K, b)$.

Then

$$T(\mu) = \mathcal{L}\left(\sum_{r=1}^{K} A_r Z^{(r)} + b\right), \quad T(\nu) = \mathcal{L}\left(\sum_{r=1}^{K} A_r W^{(r)} + b\right).$$

Proof

$$T(\mu) = \mathcal{L}\left(\sum_{r=1}^{K} A_r Z^{(r)} + b\right), \quad T(\nu) = \mathcal{L}\left(\sum_{r=1}^{K} A_r W^{(r)} + b\right).$$

Hence

$$\begin{aligned}
\ell_p(T(\mu), T(\nu)) &= \inf\{\|X - Y\|_p : \mathcal{L}(X) = T(\mu), \mathcal{L}(Y) = T(\nu)\} \\
&\leq \left\| \sum_{r=1}^{K} A_r Z^{(r)} + b - \left(\sum_{r=1}^{K} A_r W^{(r)} + b\right) \right\|_p \\
&= \left\| \sum_{r=1}^{K} A_r (Z^{(r)} - W^{(r)}) \right\|_p \leq \sum_{r=1}^{K} \|A_r\|_p \left\| Z^{(r)} - W^{(r)} \right\|_p \\
&= \left(\sum_{r=1}^{K} \|A_r\|_p\right) \left\| Z^{(1)} - W^{(1)} \right\|_p \\
&= \left(\sum_{r=1}^{K} \|A_r\|_p\right) \ell_p(\mu, \nu). \quad \clubsuit
\end{aligned}$$

Lipschitz on $(\mathcal{M}_2(0), \ell_2)$

Theorem: Assume $(A_1, \ldots, A_k, b)$ are L^2-integr. r.v. with $\mathbb{E}\, b = 0$,

$$\begin{aligned} T : \mathcal{M}_2(0) &\to \mathcal{M}_2(0) \\ \mu &\mapsto \mathcal{L}\left(\sum_{r=1}^{K} A_r Z^{(r)} + b \right), \end{aligned}$$

where $(A_1, \ldots, A_k, b), Z^{(1)}, \ldots, Z^{(K)}$ are indep. and $\mathcal{L}(Z^{(r)}) = \mu$.

Then, for all $\mu, \nu \in \mathcal{M}_2(0)$,

$$\ell_2(T(\mu), T(\nu)) \leq \left(\sum_{r=1}^{K} \mathbb{E}\, A_r^2 \right)^{1/2} \ell_2(\mu, \nu).$$

If $\sum_{r=1}^{K} \mathbb{E}\, A_r^2 < 1$ then T is a contraction on $(\mathcal{M}_2(0), \ell_2)$.

Convergence: Quickselect I

$$Y_n \stackrel{d}{=} \frac{I_n}{n} Y_{I_n} + \frac{n-1}{n}, \qquad \mathcal{L}(I_n) = \mathrm{unif}\{0, \ldots, n-1\}$$

$$Y \stackrel{d}{=} UY + 1, \qquad \mathcal{L}(U) = \mathrm{unif}[0,1].$$

We choose $I_n = \lfloor nU \rfloor$ and (Y_j, Y) as optimal couplings. Then

$$\begin{aligned}
\Delta(n) &:= \ell_p(\mathcal{L}(Y_n), \mathcal{L}(Y)) \leq \left\| \frac{I_n}{n} Y_{I_n} + \frac{n-1}{n} - (UY+1) \right\|_p \\
&\leq \left\| \frac{I_n}{n} Y_{I_n} \pm \frac{I_n}{n} Y - UY \right\|_p + \frac{1}{n} \\
&\leq \left\| \frac{I_n}{n} \left(Y_{I_n} - Y \right) \right\|_p + \frac{1 + \|Y\|_p}{n}.
\end{aligned}$$

Convergence: Quickselect II

$$\Delta(n) \le \left\| \frac{I_n}{n} \left(Y_{I_n} - Y\right) \right\|_p + \frac{1 + \|Y\|_p}{n}.$$

Now,

$$\left\| \frac{I_n}{n} \left(Y_{I_n} - Y\right) \right\|_p^p = \mathbb{E}\left[\left| \frac{I_n}{n} \left(Y_{I_n} - Y\right) \right|^p \right] = \frac{1}{n} \sum_{j=0}^{n-1} \left(\frac{j}{n}\right)^p \mathbb{E}\left[|Y_j - Y|^p\right]$$

$$= \frac{1}{n} \sum_{j=0}^{n-1} \left(\frac{j}{n}\right)^p \ell_p^p(\mathcal{L}(Y_j), \mathcal{L}(Y)) = \left\| \frac{I_n}{n} \Delta(I_n) \right\|_p^p.$$

Together

$$\Delta(n) \le \left\| \frac{I_n}{n} \Delta(I_n) \right\|_p + \frac{1 + \|Y\|_p}{n}.$$

We obtain $\Delta(n) \to 0$.

(E.g., for $p = 1$ show $\Delta(n) \le (C \log n)/n$ by induction.)

General theorem in $\mathcal{M}_p$

Let $(Y_n)_{n\geq 0}$ be L^p-integrable, $p \geq 1$, with (as before)

$$Y_n \stackrel{d}{=} \sum_{r=1}^{K} A_r^{(n)} Y^{(r)}_{I_r^{(n)}} + b^{(n)}.$$

Assume that

$$\left(A_1^{(n)}, \ldots, A_K^{(n)}, b^{(n)}\right) \xrightarrow{\ell_p} (A_1^*, \ldots, A_K^*, b^*),$$

$$\sum_{r=1}^{K} \|A_r^*\|_p < 1,$$

$$\mathbb{E}\left[\mathbf{1}_{\{I_r^{(n)} \leq \ell\}} \left|A_r^{(n)}\right|^p\right] \xrightarrow{n\to\infty} 0, \quad \forall\, \ell \in \mathbb{N}, r = 1, \ldots, K.$$

Then

$$\ell_p(\mathcal{L}(Y_n), \mathcal{L}(Y)) \xrightarrow{n\to\infty} 0,$$

where $\mathcal{L}(Y)$ is the fixed-point of T in $\mathcal{M}_p$.

General theorem in $\mathcal{M}_2(0)$

Let $(Y_n)_{n\geq 0}$ be L^2-integrable, with $\mathbb{E}\, Y_n = 0$ for all $n \geq 0$ and (as before)

$$Y_n \stackrel{d}{=} \sum_{r=1}^{K} A_r^{(n)} Y_{I_r^{(n)}}^{(r)} + b^{(n)}.$$

Assume that

$$\left(A_1^{(n)}, \dots, A_K^{(n)}, b^{(n)}\right) \xrightarrow{\ell_2} (A_1^*, \dots, A_K^*, b^*)\,,$$

$$\sum_{r=1}^{K} \mathbb{E}\left[(A_r^*)^2\right] < 1,$$

$$\mathbb{E}\left[\mathbf{1}_{\{I_r^{(n)} \leq \ell\}} \left|A_r^{(n)}\right|^2\right] \xrightarrow{n\to\infty} 0, \quad \forall\, \ell \in \mathbb{N}, r = 1, \dots, K.$$

Then

$$\ell_2(\mathcal{L}(Y_n), \mathcal{L}(Y)) \xrightarrow{n\to\infty} 0,$$

where $\mathcal{L}(Y)$ is the fixed-point of T in $\mathcal{M}_2(0)$.

Application: Path length in BST

$$Y_n \stackrel{d}{=} \frac{I_n}{n} Y^*_{I_n} + \frac{n-1-I_n}{n} Y^{**}_{n-1-I_n} + b^{(n)},$$

$$Y \stackrel{d}{=} UY^* + (1-U)Y^{**} + g(U).$$

Convergence of coefficients and technical condition satisfied.

Contraction: $A_1^* = U$, $A_2^* = 1 - U$.

In $(\mathcal{M}_p, \ell_p)$: NO for all $p \geq 1$:

$$\|A_1^*\|_p + \|A_2^*\|_p \geq \|A_1^*\|_1 + \|A_2^*\|_1 = \mathbb{E}\,U + \mathbb{E}\,(1-U) = 1.$$

In $(\mathcal{M}_2(0), \ell_2)$: YES:

$$\mathbb{E}\left[(A_1^*)^2\right] + \mathbb{E}\left[(A_2^*)^2\right] = \mathbb{E}\,U^2 + \mathbb{E}\,(1-U)^2 = \frac{2}{3} < 1.$$

Application: Central limit theorem

Let $W_1, W_2, \ldots$ be i.i.d., L^p-integrable, $p \geq 2$,
with $\mathbb{E}\, W_1 = \mu$, $\mathrm{Var}(W_1) = \sigma^2$.

Consider

$$X_n := \sum_{i=1}^{n} W_i \stackrel{d}{=} X^*_{\lceil n/2 \rceil} + X^{**}_{\lfloor n/2 \rfloor}.$$

Then we have

$$Y_n := \frac{X_n - \mathbb{E}\, X_n}{\sqrt{\mathrm{Var}(X_n)}} = \frac{X_n - \mu n}{\sigma \sqrt{n}}$$

$$\stackrel{d}{=} \sqrt{\frac{\lceil n/2 \rceil}{n}}\, Y^*_{\lceil n/2 \rceil} + \sqrt{\frac{\lfloor n/2 \rfloor}{n}}\, Y^{**}_{\lfloor n/2 \rfloor}.$$

Limit equation:

$$Y \stackrel{d}{=} \frac{1}{\sqrt{2}} Y^* + \frac{1}{\sqrt{2}} Y^{**}.$$

Contraction properties

$$Y \stackrel{d}{=} \frac{1}{\sqrt{2}} Y^* + \frac{1}{\sqrt{2}} Y^{**}.$$

Contraction: $A_1^* = A_2^* = 1/\sqrt{2}$

In $(\mathcal{M}_p, \ell_p)$: NO for all $p \geq 1$: $\|A_1^*\|_p + \|A_2^*\|_p = \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} = \sqrt{2} > 1.$

In $(\mathcal{M}_2(0), \ell_2)$: NO: $\mathbb{E}\left[(A_1^*)^2\right] + \mathbb{E}\left[(A_2^*)^2\right] = \frac{1}{2} + \frac{1}{2} = 1.$

Same problem for

$$Y \stackrel{d}{=} \sum_{r=1}^{K} A_r^* Y^{(r)},$$

with random $(A_1^*, \ldots, A_K^*)$ with

$$\sum_{r=1}^{K} (A_r^*)^2 = 1 \qquad \text{almost surely.}$$

The fundamental problem

$$Y \stackrel{d}{=} \sum_{r=1}^{K} A_r^* Y^{(r)}, \qquad \sum_{r=1}^{K} (A_r^*)^2 = 1 \quad \text{a.s.}$$

Appears in CLT, space of m-ary search tree $(3 \leq m \leq 26)$, number of maxima in right triangles, etc.

Easy to check:

$$\mathcal{N}(0, \sigma^2) \text{ is solution for all } \sigma \geq 0.$$

However

$$\mathcal{N}(0, \sigma^2) \in \mathcal{M}_2(0) \text{ for all } \sigma \geq 0.$$

Also $\mathcal{N}(0, \sigma^2) \in \mathcal{M}_p$ for all $p \geq 1$, $\sigma \geq 0$.

$\Longrightarrow$ no unique fixed-point in these spaces.

$\Longrightarrow$ no contraction in any metric in these spaces.

Zolotarev metric

$$\zeta_s(\mathcal{L}(X), \mathcal{L}(Y)) = \sup_{f \in \mathcal{F}_s} |\mathbb{E}[f(X) - f(Y)]|, \quad s > 0,$$

with

$$\mathcal{F}_s = \{f \in C^m(\mathbb{R}, \mathbb{R}) : |f^{(m)}(x) - f^{(m)}(y)| \leq |x - y|^\alpha\},$$

$$s = m + \alpha, \quad m \in \mathbb{N}_0, \; \alpha \in (0, 1].$$

Functions in $\mathcal{F}_s$:

$$x \mapsto \frac{1}{2|t|^s} \cos(tx), \; t \neq 0; \quad x \mapsto \frac{1}{s(s-1)\cdots(s-m+1)} |x|^s$$

For $f \in \mathcal{F}_s$: $\quad x \mapsto f(x+z), \; z \in \mathbb{R}; \quad x \mapsto \frac{1}{|c|^s} f(cx), \; c \neq 0.$

Zolotarev metric II

Spaces of probability measures

$$\mathcal{M}_s(M_1,\dots,M_m) := \left\{\mu \in \mathcal{M}_s \,\middle|\, \int x^k d\mu(x) = M_k, k = 1,\dots,m\right\}$$
$$= \left\{\mathcal{L}(X) \in \mathcal{M}_s \,\middle|\, \mathbb{E}\left[X^k\right] = M_k, k = 1,\dots,m\right\}.$$

Lemma: ζ_s, $s > 0$, is finite on $\mathcal{M}_s(M_1,\dots,M_m)\times\mathcal{M}_s(M_1,\dots,M_m)$.
Furthermore, we have

$$\zeta_s(\mathcal{L}(X),\mathcal{L}(Y)) \le \frac{\Gamma(1+\alpha)}{\Gamma(1+s)} \left(\mathbb{E}\,|X|^s + \mathbb{E}\,|Y|^s\right).$$

Proof.
Exercise. ♣

Short Notation: $\zeta_s(X,Y) := \zeta_s(\mathcal{L}(X),\mathcal{L}(Y))$.

Zolotarev metric: properties

Lemma:

a) ζ_s is $(s,+)$-ideal, i.e.,
for $c \neq 0$ and Z independent of (X, Y) we have

$$\zeta_s(X+Z, Y+Z) \leq \zeta_s(X,Y), \qquad \zeta_s(cX, cY) = |c|^s \zeta_s(X,Y).$$

b) If $\zeta_s(X_n, X) \to 0$, then $X_n \xrightarrow{d} X$ and $\|X_n\|_s \to \|X\|_s$.

c) If $X_1, \ldots, X_K, Y_1, \ldots, Y_K$ independent, then

$$\zeta_s\left(\sum_{r=1}^{K} X_r, \sum_{r=1}^{K} Y_r\right) \leq \sum_{r=1}^{K} \zeta_s(X_r, Y_r).$$

Zolotarev metric: properties

Proof.

a) Note that $x \mapsto f(x+z)$ in $\mathcal{F}_s$ for $f \in \mathcal{F}_s$.

$$\begin{aligned}
\zeta_s(X+Z, Y+Z) &= \sup_{f\in\mathcal{F}_s} \left|\mathbb{E}\left[f(X+Z) - f(Y+Z)\right]\right| \\
&= \sup_{f\in\mathcal{F}_s} \left|\int \mathbb{E}\left[f(X+z) - f(Y+z)\right] d\mathbb{P}_Z(z)\right| \\
&\leq \int \underbrace{\sup_{f\in\mathcal{F}_s} \left|\mathbb{E}\left[f(X+z) - f(Y+z)\right]\right|}_{=\sup_{f\in\mathcal{F}_s}|\mathbb{E}[f(X)-f(Y)]| = \zeta_s(X,Y)} d\mathbb{P}_Z(z) \\
&= \zeta_s(X, Y).
\end{aligned}$$

Second part similar with $x \mapsto |c|^{-s} f(cx)$ in $\mathcal{F}_s$ for $f \in \mathcal{F}_s$. ♣

Zolotarev metric: properties

Proof.

b) Characteristic functions:

$$\varphi_{X_n}(t) = \mathbb{E}\exp(itX_n) = \mathbb{E}\cos(tX_n) + i\mathbb{E}\sin(tX_n).$$

Now, $x \mapsto \frac{1}{2|t|^s}\cos(tx)$ is in $\mathcal{F}_s$.

Hence, $\zeta_s(X_n, X) \to 0$ implies $|\varphi_{X_n}(t) - \varphi_X(t)| \to 0$ for all $t \in \mathbb{R}$.

Continuity theorem of Lévy implies $X_n \xrightarrow{d} X$.

Second assertion use $x \mapsto \frac{1}{s(s-1)\cdots(s-m+1)}|x|^s$ in $\mathcal{F}_s$. ♣

c) Exercise (easy). ♣

Lipschitz in ζ_s

Theorem: Assume $(A_1, \ldots, A_k, b)$ are L^s-integrable r.v. and

$$\begin{aligned} T : \mathcal{M} &\to \mathcal{M} \\ \mu &\mapsto \mathcal{L}\left(\sum_{r=1}^{K} A_r Z^{(r)} + b \right), \end{aligned}$$

where $(A_1, \ldots, A_k, b), Z^{(1)}, \ldots, Z^{(K)}$ are indep. and $\mathcal{L}(Z^{(r)}) = \mu$.

Then, for all $\mu, \nu \in \mathcal{M}_s(M_1, \ldots, M_m)$,

$$\zeta_s(T(\mu), T(\nu)) \leq \left(\sum_{r=1}^{K} \mathbb{E}\left[|A_r|^s \right] \right) \zeta_s(\mu, \nu).$$

Lipschitz in ζ_s

Proof.
$Z^{(1)}, \dots, Z^{(K)}, W^{(1)}, \dots, W^{(K)}, (A_1, \dots, A_k, b)$ independent with $\mathcal{L}(Z^{(r)}) = \mu,\ \mathcal{L}(W^{(r)}) = \nu$.

Hence

$$T(\mu) = \mathcal{L}\left(\sum_{r=1}^{K} A_r Z^{(r)} + b\right), \quad T(\nu) = \mathcal{L}\left(\sum_{r=1}^{K} A_r W^{(r)} + b\right).$$

Then

$$\zeta_s(T(\mu), T(\nu)) = \sup_{f \in \mathcal{F}_s} \left| \mathbb{E}\left[f\left(\sum_{r=1}^{K} A_r Z^{(r)} + b\right) - f\left(\sum_{r=1}^{K} A_r W^{(r)} + b\right)\right]\right|$$

Lipschitz in ζ_s

Denote $\sigma = \mathbb{P}_{(A_1,\ldots,A_K,b)}$ and $\mathbf{x} = (\alpha_1, \ldots, \alpha_K, \beta)$. Then

$$
\begin{aligned}
\zeta_s(T(\mu), T(\nu)) &= \sup_{f\in\mathcal{F}_s} \left| \mathbb{E}\left[f\left(\sum_{r=1}^K A_r Z^{(r)} + b\right) - f\left(\sum_{r=1}^K A_r W^{(r)} + b\right)\right]\right| \\
&= \sup_{f\in\mathcal{F}_s} \left| \int \mathbb{E}\left[f\left(\sum_{r=1}^K \alpha_r Z^{(r)} + \beta\right) - f\left(\sum_{r=1}^K \alpha_r W^{(r)} + \beta\right)\right] d\sigma(\mathbf{x}) \right| \\
&\le \int \sup_{f\in\mathcal{F}_s} \left| \mathbb{E}\left[f\left(\sum_{r=1}^K \alpha_r Z^{(r)} + \beta\right) - f\left(\sum_{r=1}^K \alpha_r W^{(r)} + \beta\right)\right]\right| d\sigma(\mathbf{x}) \\
&\le \int \zeta_s\left(\sum_{r=1}^K \alpha_r W^{(r)} + \beta, \sum_{r=1}^K \alpha_r Z^{(r)} + \beta\right) d\sigma(\mathbf{x}) \\
&\overset{a),c)}{\le} \int \sum_{r=1}^K \zeta_s\left(\alpha_r W^{(r)}, \alpha_r Z^{(r)}\right) d\sigma(\mathbf{x}) \overset{a)}{\le} \int \sum_{r=1}^K |\alpha_r|^s \zeta_s\left(W^{(r)}, Z^{(r)}\right) d\sigma(\mathbf{x}) \\
&= \left(\sum_{r=1}^K \mathbb{E}\left[|A_r|^s\right]\right) \zeta_s(\mu, \nu). \quad \clubsuit.
\end{aligned}
$$

The fundamental problem II

$$T : \mathcal{M} \to \mathcal{M}, \quad \mu \mapsto \mathcal{L}\left(\sum_{r=1}^{K} A_r^* Z^{(r)}\right), \qquad \text{with } \sum_{r=1}^{K} (A_r^*)^2 = 1 \quad \text{a.s.,}$$

where $(A_1^*, \ldots, A_K^*), Z^{(1)}, \ldots, Z^{(K)}$ independent and $\mathcal{L}(Z^{(r)}) = \mu$.

Excluded case: $\mathbb{P}\left(\bigcup_{r=1}^{K} \{A_r^* = 1\}\right) = 1.$ $(*)$

Known: $\{\mathcal{N}(0, \sigma^2) \mid \sigma \geq 0\}$ are fixed-points.

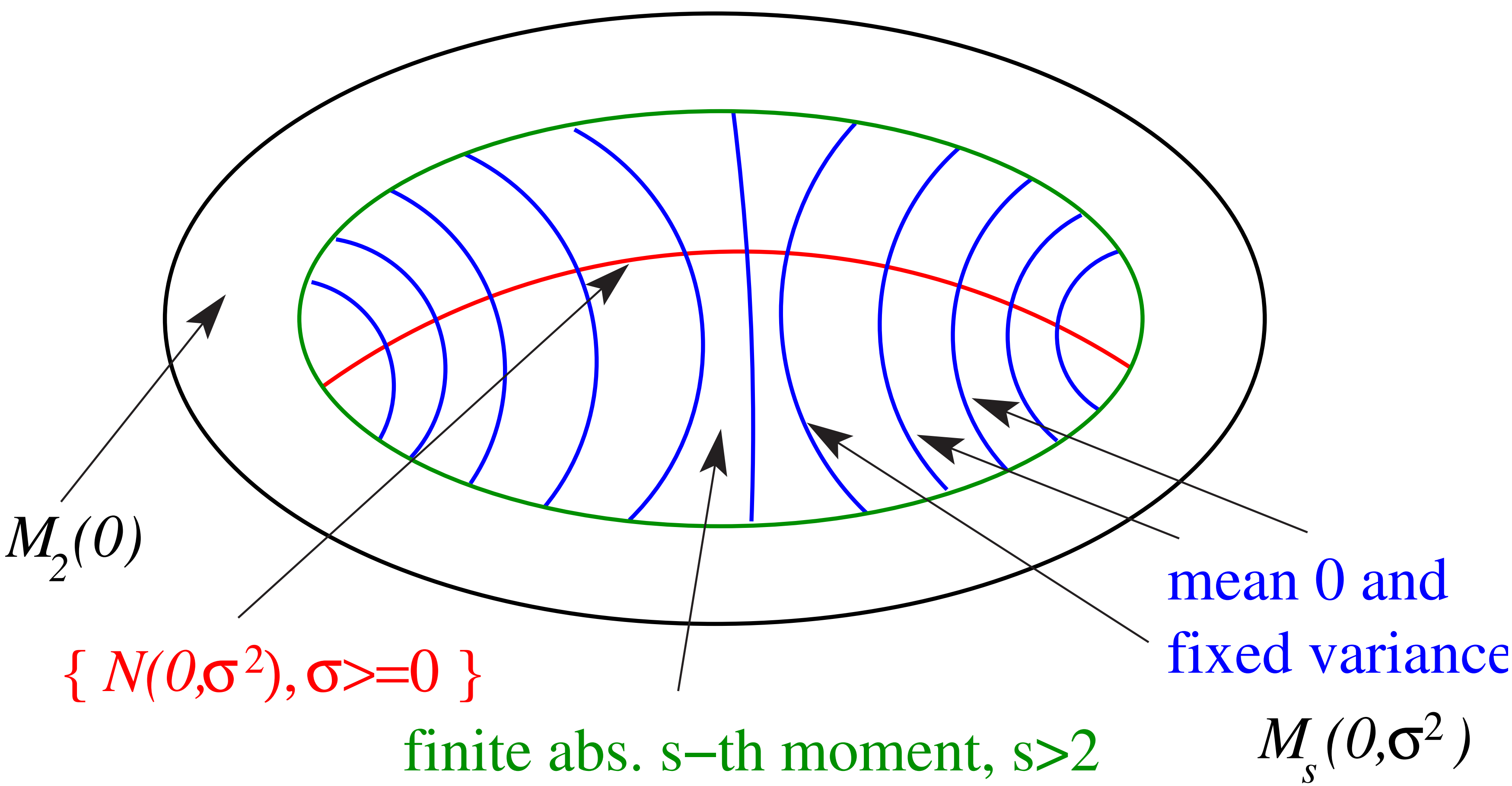
We have $T(\mathcal{M}_s(0,1)) \subseteq \mathcal{M}_s(0,1)$ for $2 < s \leq 3$.

$(*)$ implies $\sum_{r=1}^{K} \mathbb{E}\left[(A_r^*)^s\right] < 1$.

Previous Lemma: T contraction on $(\mathcal{M}_s(0,1), \zeta_s)$, $2 < s \leq 3$.

$\Rightarrow \mathcal{N}(0,1)$ unique fixed point of T in $\mathcal{M}_s(0,1)$.

The work space



Central limit theorem II

$W_1, W_2, \ldots$ i.i.d., L^s-integr., $s > 2$, with $\mathbb{E}\, W_1 = \mu$, $\mathrm{Var}(W_1) = \sigma^2$.

$$X_n \; := \; \sum_{i=1}^{n} W_i \overset{d}{=} X^*_{\lceil n/2 \rceil} + X^{**}_{\lfloor n/2 \rfloor},$$

$$Y_n := \frac{X_n - \mu n}{\sigma \sqrt{n}} \; \overset{d}{=} \; \sqrt{\frac{\lceil n/2 \rceil}{n}} Y^*_{\lceil n/2 \rceil} + \sqrt{\frac{\lfloor n/2 \rfloor}{n}} Y^{**}_{\lfloor n/2 \rfloor}.$$

Consider $2 < s \leq 3$. For $\mathcal{L}(N) = \mathcal{N}(0,1)$ we have

$$\mathbb{E}\, Y_n^k = \mathbb{E}\, N^k \text{ for } k = 1, 2 \quad \text{and} \quad \mathbb{E}\, |Y_n^k|^s, \mathbb{E}\, |N|^s < \infty.$$

$$\Rightarrow \zeta_s(Y_n, N) < \infty \text{ for all } n \geq 1.$$

Convergence

$$
\begin{aligned}
Y_n &\stackrel{d}{=} \sqrt{\frac{\lceil n/2\rceil}{n}} Y^*_{\lceil n/2\rceil} + \sqrt{\frac{\lfloor n/2\rfloor}{n}} Y^{**}_{\lfloor n/2\rfloor}, \\
N &\stackrel{d}{=} \sqrt{\frac{\lceil n/2\rceil}{n}} N^* + \sqrt{\frac{\lfloor n/2\rfloor}{n}} N^{**}.
\end{aligned}
$$

Thus

$$
\begin{aligned}
\zeta_s(Y_n, N) &= \zeta_s\left(\sqrt{\frac{\lceil n/2\rceil}{n}} Y^*_{\lceil n/2\rceil} + \sqrt{\frac{\lfloor n/2\rfloor}{n}} Y^{**}_{\lfloor n/2\rfloor}, \sqrt{\frac{\lceil n/2\rceil}{n}} N^* + \sqrt{\frac{\lfloor n/2\rfloor}{n}} N^{**}\right) \\
&\stackrel{c)}{\leq} \zeta_s\left(\sqrt{\frac{\lceil n/2\rceil}{n}} Y^*_{\lceil n/2\rceil}, \sqrt{\frac{\lceil n/2\rceil}{n}} N^*\right) + \zeta_s\left(\sqrt{\frac{\lfloor n/2\rfloor}{n}} Y^{**}_{\lfloor n/2\rfloor}, \sqrt{\frac{\lfloor n/2\rfloor}{n}} N^{**}\right) \\
&\stackrel{a)}{\leq} \left(\frac{\lceil n/2\rceil}{n}\right)^{s/2} \zeta_s\left(Y^*_{\lceil n/2\rceil}, N^*\right) + \left(\frac{\lfloor n/2\rfloor}{n}\right)^{s/2} \zeta_s\left(Y^{**}_{\lfloor n/2\rfloor}, N^{**}\right)
\end{aligned}
$$

Convergence II

With $\Delta(n) := \zeta_s(Y_n, N)$ we obtain

$$\Delta(n) \leq \left(\frac{\lceil n/2 \rceil}{n}\right)^{s/2} \Delta\left(\lceil n/2 \rceil\right) + \left(\frac{\lfloor n/2 \rfloor}{n}\right)^{s/2} \Delta\left(\lfloor n/2 \rfloor\right).$$

Easy: $\Delta(n) \leq \zeta_s(Y_1, N)$ for all $n \geq 1$ (by induction).

With $L := \limsup_{n\to\infty} \Delta(n)$ we obtain

$$L \leq 2^{1-s/2} L, \qquad \text{hence} \qquad L = 0.$$

A useful extension

Theorem:

$$X_n \stackrel{d}{=} \sum_{r=1}^{K} X^{(r)}_{I_r^{(n)}} + b_n, \quad n > n_0$$

with conditions as before and L^s-integrable, $2 < s \leq 3$. Assume

$$\mathbb{E}\, X_n = \mu(n) + o(\sigma(n)), \quad \mathrm{Var}(X_n) = \sigma^2(n) + o(\sigma^2(n))$$

and

$$\left(\frac{\sigma(I_1^{(n)})}{\sigma(n)}, \dots, \frac{\sigma(I_K^{(n)})}{\sigma(n)}\right) \xrightarrow{\ell_s} (A_1^*, \dots, A_K^*), \quad \sum_{r=1}^{K} (A_r^*)^2 = 1 \text{ a.s. with } (*),$$

$$b_n + \mu(n) - \sum_{r=1}^{K} \mu(I_r^{(n)}) = o(\sigma(n)) \text{ in } L_s.$$

Then

$$\frac{X_n - \mathbb{E}\, X_n}{\sqrt{\mathrm{Var}(X_n)}} \xrightarrow{d} \mathcal{N}(0, 1).$$