Talagrand's simplex construction

The following note is based on Talagrand (2021, §2.6)

1 A gaussian process

S:gaussian

Let T be the unit simplex in $\mathbb{R}^M$, with $M \geq 2$ to avoid total triviality. That is, each member of T is an element of $\mathbb{R}^M$ for which each coordinate t_α (also written as $t[\alpha]$ to avoid double subscripting) is nonnegative and $\sum_{\alpha=1}^M t_\alpha = 1$. Equivalently, T equals the convex hull of the usual orthonomal basis, $\{e_\alpha : 1 \leq \alpha \leq M\}$. Those vectors constitute the extreme points of the compact, convex set T.

For the purposes of this note it helps to have an ordering imposed on the coordinates. That is, each M-tuple of real numbers is explicitly thought of as a map from the ordered set $[[M]] := \{\alpha \in \mathbb{N} : 1 \leq \alpha \leq M\}$ into $\mathbb{R}$. To emphasize the ordering I'll often write $\mathbb{R}^{[[M]]}$ instead of $\mathbb{R}^M$.

To simplify notation slightly, I'll assume that $M = 2^m$ for some positive integer m. Results for general M can easily be deduced from the results for the m defined by $2^{m-1} < M \leq 2^m$.

Let $g := \{g_\alpha : \alpha \in [[M]]\}$ be a vector of independent standard normal variates. The process defined by

$$X_t := \sum\nolimits_\alpha t_\alpha g_\alpha = \langle t, g\rangle \qquad \text{for } t \in T$$

is centered gaussian with

$$\operatorname{cov}(X_s, X_t) = \sum\nolimits_\alpha s_\alpha t_\alpha = \langle s, t\rangle \qquad \text{for } s, t \in T.$$

I'll write $|s|_2$ for $\sqrt{\langle s, s\rangle}$, the usual euclidean norm.

As Talagrand noted, $\sup_{t\in T} X_t = \max_\alpha g_\alpha$ because the supremum of a linear function on the compact, convex set T is achieved at one of its extreme points. It follows that

$$\mathbb{P}\sup_{t\in T} X_t \le C\sqrt{\log M}$$

for some constant C that doesn't depend on M. He then sought to construct a lower bound using one of the majorizing measure equivalences, a task that he noted to be non-trivial.

More precisely, Talagrand showed how to construct subsets $T_n \subset \mathbb{R}_+^{[[M]]}$ with

\E@ Tal2.6.1 <1>
$$T_0 = \{0\} \text{ and } |T_n| \le N_n := 2^{2^n} \text{ and}$$
$$\sup_{t\in T}\sum\nolimits_{n\ge 0} 2^{n/2} d(t, T_n) \le C_1\sqrt{\log M} = C_2\sqrt{m} \qquad \text{if } M = 2^m,$$

for some constants C_1 and C_2 that also don't depend on M.

The subset T_n could also be replaced by $\{\pi(s) : s \in T_n\}$, where π denotes the map that takes each s in $\mathbb{R}^M$ to its closest point in T. For each s in T_n and t in T we have $|s-t|_2 \ge |s-\pi(s)|_2$, which implies

$$|t-\pi(s)|_2 \le |t-s|_2 + |s-\pi(s)|_2 \le 2|t-s|_2.$$

2 The main idea

S:idea

The basic approximation method starts with a vector s in $\mathbb{R}_+^{[[M]]}$ and a positive integer p. The operation splits s into a sum $A+R$. The coordinates for the subset $I := \{\alpha \in [[M]] : s_\alpha > 2^{-p}\}$ are reduced by some multiple of 2^{-p}; the remaining coordinates are left untouched: $A_\alpha = 0$ and $R_\alpha = s_\alpha$ for $\alpha \notin I$. For α in I,

$$A_\alpha = \lambda_\alpha 2^{-p} \text{ where } \lambda_\alpha := \lfloor 2^p s_\alpha \rfloor \quad \text{AND} \quad R_\alpha = s_\alpha - A_\alpha.$$

The method ensures that $0 \le R_\alpha \le 2^{-p}$ for all α in $[[M]]$. It also results in smallish λ_α's if each s_α is bounded above by a small muliple of 2^{-p}. For example, if $s_\alpha \le 4\times 2^{-p}$ for each α then $\lambda_\alpha \in \{0,1,2,3,4\}$. The size of this set of multiples together with the size of I controls the cardinality of the set of all possible approximating A vectors.

> **Remark.** Talagrand actually used the largest integer multiple of 2^{-p} that is *strictly smaller* than s_α. For example, if $s_\alpha = 4\times 2^{-p}$ he would use $\lambda_\alpha = 3$, whereas my definition uses $\lambda_\alpha = 4$. I don't think the difference matters: later in the argument I get 5 where he used the constant 4.

The desire for control over $\max_\alpha s_\alpha$ suggests a recursive argument. To each t in T we must construct an increasing sequence of integers

$$p_0 = 0 \le p_1 = p[1,t] \le p_2 = p[2,t] \le \dots$$

with $p_n \le 2 + p_{n-1}$ for $n \in \mathbb{N}$ to ensure that $2^{-p[n,t]} \le 4 \times 2^{-p[n-1,t]}$. Initially $t_\alpha \le 1 = 2^0$ for every α. Using p_1 we split t into $A^{(1)} + R^{(1)}$ with $\max_\alpha R^{(1)}_\alpha \le 2^{-p[1,t]}$. The coordinate $A^{(1)}_\alpha$ will equal 0 for each α not in the set $I_1(t) := \{\alpha : t_\alpha > 2^{-p[1,t]}\}$ and

$$A^{(1)}_\alpha = \lambda^{(1)}_\alpha 2^{-p[1,t]} \qquad \text{with } \lambda^{(1)}_\alpha \in \{0,1,2,3,4\} \text{ if } \alpha \in I_1(t)$$

Using the same $\{p_n\}$ sequence we then split $R^{(1)}$ into $A^{(2)} + R^{(2)}$ with $\max_\alpha R^{(2)} \le 2^{-p[2,t]}$. And so on. The vector $\tau^{(n)}(t) := \sum_{k=1}^{n} A^{(k)}$ becomes the nth approximation to t. As t ranges over T we will get a large collection of approximating vectors, $\{\tau^{(n)}(t) : t \in T\}$. Control over the size of that set will enable us to bound the size of the approximating set T_n in <1>.

3 The $p[n,t]$ sequence

S:pn

The main ingredient for the argument sketched in Section 2 is the sequence of integers $p_n = p[n,t]$ for each t in T, with the properties stated in the following Lemma.

Tal2.6.2 <2> **Lemma.** *For each t in the simplex T there exists an increasing integers $\{p[n,t] : n \in \mathbb{N}_0\}$ for which:*

(i) $0 = p[0,t] \le p[n,t] \le 2n$ *for each n.*

(ii) $p[n+1,t] \le 2 + p[n,t]$ *for each n.*

(iii) The set $\mathcal{H}_n(t) := \{\alpha \in [[M]] : t_\alpha > 2^{-p[n,t]}\}$ *has size* $< 2^n$ *for each n.*

(iv) $\sum_\alpha t_\alpha^2 \{2^{-p[n-1,t]} \ge t_\alpha > 2^{-2p[n,t]}\} \le 2^{n-p[n-1,t]}$ *for each n.*

□ *(v)* $\sup_{t\in T} \sum_{n\in\mathbb{N}} 2^{n-p[n,t]} \le C_2$ *for some C_2 not depending on M.*

Remark. Notice that

$$\emptyset = \mathcal{H}_0(t) \subset \mathcal{H}_1(t) \subset \cdots \subset \mathcal{H}_n(t) \uparrow \{\alpha \in [[M]] : t_\alpha > 0\}$$

and $\{t \in T : 2^{-p[n-1,t]} \ge t_\alpha > 2^{-2p[n,t]}\} = \mathcal{H}_{n-1}(t)^c \cap \mathcal{H}_n(t)$.

The rest of this Section proves the Lemma, following some general comments about Talagrand's approach.

The first thing to note is that the p_n's for a given t will depend only on the M-vector s obtained by sorting the $t[\alpha] := t_\alpha$ coordinates into decreasing order. To avoid minor notational complications caused by finiteness of $[[M]]$ it helps to embed s in an infinite sequence by defining $s_\alpha := 0$ for $\alpha > M$. In fact the construction depends on s only through the values $s[2^k]$ for $k \in [[m]]$. The underlying reason for this simplification is revealed by partitioning the index set $[[M]]$ into disjoint blocks $B_0 = \{1\}$, $B_1 = \{2\}$, $B_3 = \{3,4\}$, and so on. That is,

$$B_k = \{\alpha \in [[M]] : 2^{k-1} < a \le 2^k\} \qquad \text{for } k \in [[m]].$$

By monotonicity,

$$1 = \sum\nolimits_{\alpha\in[\![M]\!]} s_\alpha = s_1 + \sum\nolimits_{k=1}^{m} \sum\nolimits_{\alpha\in B_k} s_\alpha \ge s_1 + \sum\nolimits_{k=1}^{m} 2^{k-1} s[2^k].$$

Thus

\E@ 2^ks <3>

$$\sum\nolimits_{k=1}^{m} 2^k s[2^k] \le 2 \sum\nolimits_{k=1}^{m} s_\alpha \le 2.$$

It will also be helpful to remember that $s[2^k] \le \sum_{\alpha=1}^{2^k} s_\alpha/2^k \le 2^{-k}$.

Remark. Talagrand included the s_1 contribution in the sum, resulting in the bound $\sum_{k=0}^{m} 2^k s[2^k] \le 3$.

At this point Talgrand defined p_n by a two-step method that seems strange to me. (I note that a similar construction appears in his §2.14, which I have not yet read carefully.) For each nonegative integer k he defined q_k as the largest integer $\le 2k$ for which $2^{-k} > s[2^k]$ and then defined

$$p_n := \min\nolimits_{0\le k\le n}\left(q_k + 2(n-k)\right).$$

I found that his argument can be slightly simplified.

Proof (of the Lemma). Working with s, the montone rearrangement of t, define

$$\beta_n := \beta[n,t] := \sup\{k \in \mathbb{N} : 2^{-k} \ge s[2^n]\} \qquad \text{for } k \in \mathbb{N}.$$

Notice that $s[2^1] \le 1/2$ because $1 \ge s[1] + s[2]$, which ensures that β_1 is well defined. Notice also that $\beta_n = +\infty$ iff $s[2^n] = 0$. In particular, we must have $\beta_n = +\infty$ if $n > m$.

Starting from $p_0 := p[0,t] := 0$, recursively define

\E@ pn.def <4>

$$p_n := p[n,t] := \min\left(2 + p[n-1,t]\,,\, \beta[n,t]\right) \qquad \text{for } n \in \mathbb{N}.$$

By construction, $p_n \in \mathbb{N}$ for $n \in \mathbb{N}$ and, because β is an increasing function,

$$p_{n+1} = \min\left(2 + p_n, \beta_{n+1}\right) \ge \min\left(p_n, \beta_n\right)) = p_n.$$

The inclusion of the β_n in the minimum ensures that $2^{-p[n,t]} \ge s[2^n]$. Moreover, if $p[n,t] = \beta_n$ then β_n is finite and

\E@ beta.finite <5>

$$2^{-p[n,t]} \ge s[2^n] > 2^{-p[n,t]-1} \qquad \text{implying} \qquad 2s[2^n] > 2^{-p[n,t]}.$$

The inclusion of the $p_{n-1}+2$ ensures that $p_n \le p_{n-1}+2$ for each n, as required by Lemma <2>(i). Part (ii) is a trivial consequence of (i). And if $s_\alpha > 2^{-p[n,t]} \ge s[2^n]$ then we must have $\alpha < 2^n$, which gives (iii).

Inequality (iv) follows from (iii) and the trivial bound $t_\alpha^2 \le 2^{-2p[n-1,t]}$ for each α in $\mathcal{H}_{n-1}(t)^c \cap \mathcal{H}_n(t)$.

The argument for (v) is more interesting. Let me temporarily write $\delta[n]$ for $n - p[n,t]$ and D_n for $2^{\delta[n]}$. For $n > m$ we have $\beta_n = +\infty$ so that $p_{n+1} = p_n + 2$ for $n \ge m$, implying

$$\delta[m+k] = m + k - (p[m] + 2k) = \delta[m] - k \qquad \text{for } k \in \mathbb{N},$$

whence

$$\sum\nolimits_{n>m} D_n = 2^{\delta[m]} \sum\nolimits_{k\in\mathbb{N}} 2^{-k} = D_m.$$

For assertion (v) of the Lemma it therefore suffices to bound $\sum_{n\in[\![M]\!]} D_n$.

For n in the set $[\![M]\!]_\beta := \{n \in [\![M]\!] : p_n = \beta_n\}$ we have $\beta_n < \infty$, so that inequality <5> gives $D_n \le 2^{n+1}s[2^n]$, implying

$$\sum\nolimits_{n\in[\![M]\!]_\beta} D_n \le \sum\nolimits_{n\in[\![M]\!]} 2^{n+1}s[2^n] \le 4.$$

If the set $[\![M]\!]\backslash[\![M]\!]_\beta$ is not empty then it consists of a union of stretches of the form $k+1, \dots k+\ell$ with either $k = 0$ or $k \in [\![M]\!]_\beta$. Within that stretch we have $p_{k+j} = p_k + 2j$ and $\delta[k+j] = \delta[k] - j$. Thus

$$\sum\nolimits_{j=1}^{\ell} D_{k+j} = D_k \sum\nolimits_{j=1}^{\ell} 2^{-j} \le D_k.$$

Summing over all such stretches we arrive at the bound

$$\sum\nolimits_{n=1}^{M} D_n \le D_0 + 2\sum\nolimits_{n\in[\![M]\!]_\beta} D_n \le 9,$$

□ the desired inequality for (v).

4 Construction of the approximations

S:construction

Now comes the recursive construction of the sequence of approximations for each t in the simplex. As in Section 2, the argument starts with $R^{(0)}(t) = t$, which it decomposes into a sum $A^{(1)}(t) + R^{(1)}(t)$ with $\max_\alpha R^{(1)}_\alpha(t) \le 2^{-p[1,t]}$. Then it decomposes $R^{(1)}(t)$ into $A^{(2)}(t) + R^{(2)}(t)$ with $\max_\alpha R^{(2)}_\alpha(t) \le 2^{-p[2,t]}$. And so on.

In general, the $R^{(n-1)}(t)$ vector (with $\max_\alpha R^{(n-1)}_\alpha(t) \le 2^{-p[n-1,t]}$) is decomposed into $A^{(n)}(t) + R^{(n)}(t)$ in the following way. Define

$$I_n(t) = \{\alpha \in [\![M]\!] : R^{(n-1)}_\alpha(t) > 2^{-p[n,t]}\}.$$

For α in $[\![M]\!]\backslash I_n(t)$ define $R^{(n)}_\alpha(t) = R^{(n-1)}_\alpha(t)$ and $A^{(n)}_\alpha(t) = 0$. For α in $I_n(t)$ define

$$A^{(n)}_\alpha(t) = \lambda_\alpha 2^{-p[n,t]} \qquad \text{where } \lambda_\alpha = \lfloor 2^{p[n,t]} R^{(n-1)}_\alpha(t) \rfloor$$

and $R^{(n)}_\alpha(t) = R^{(n-1)}_\alpha(t) - A^{(n)}_\alpha(t)$, which is $\le 2^{-p[n,t]}$ by definition of the floor operation $\lfloor \dots \rfloor$. Because

$$0 \le R^{(n-1)}_\alpha(t) \le 2^{-p[n-1,t]} \le 2^{-p[n,t]+2} \qquad \text{by Lemma <2>(ii)},$$

we must have $\lambda_\alpha \in \{0, 1, 2, 3, 4\}$.

We can think of $\tau^{(k)}(t) := \sum_{j=1}^{k} A^{(j)}(t)$ as the kth approximation to t and $t - \tau^{(n)}(k) = R^{(n)}(t)$ as the error of the approximation. The set of possible approximating values is then

$$U_k := \{\tau^{(k)}(t) : t \in T\}$$

and $d(t, U_k)$, the euclidean distance of t to U_k, is less than $|R^{(k)}(t)|_2$. We also need to control the size of U_k to derive the result <2>.

Claim 1: For all t in T and k in $\mathbb{N}$,

$$d(t, U_k) \leq \sum\nolimits_{j \in \mathbb{N}} \{j \geq k\} 2^{j/2 - p[j-1,t]}.$$

Claim 2: There is a universal constant c such that $|U_k| \leq M^{c2^k}$ if $M \geq 2$. That is, $\log_2 \log_2 |U_k| \leq k + \log_2(cm)$.

For approximation <1> we need $\log_2 \log_2 |T_n| \leq n$, with no dependence on M, which suggests we define $T_{n_m+k} = U_k$ for $k \in \mathbb{N}$, where $n_m := \lceil \log_2(cm) \rceil$. We then have the desired upper bound on $|T_n|$ for $n > n_m$. For $n \leq n_m$ we can just take T_n as the singleton set $\{0\}$. With those choices we get $\log_2 |T_n| \leq 2^n$ for all n, as required by the first part of <1>. For the second part, use **Claim 1** with $d(t, T_n) \leq 1$ for $n \leq n_m$ and $d(t, T_{n_m+k}) = d(t, U_k)$ for $k \in \mathbb{N}$ to get

$$\begin{aligned}
&\sum\nolimits_{n \geq 0} 2^{n/2} d(t, T_n) \\
&\qquad \leq \sum\nolimits_{n=0}^{n_m} 2^{n/2} + \sum\nolimits_{k \in \mathbb{N}} 2^{(n_m+k)/2} \sum\nolimits_{j \in \mathbb{N}} \{j \geq k\} 2^{j/2 - p[j-1,t]} \\
&\qquad \leq 2^{1+n_m/2} + \sum\nolimits_{j \in \mathbb{N}} 2^{j/2 - p[j-1,t]} \sum\nolimits_{k \in \mathbb{N}} \{j \geq k\} 2^{(n_m+k)/2} \\
&\qquad \leq 2^{n_m/2} \left(2 + \sum\nolimits_{j \in \mathbb{N}} 2^{j/2 - p[j-1,t]} 2^{(j+1)/2} / (\sqrt{2} - 1) \right).
\end{aligned}$$

The factor $2^{n_m/2}$ is $\leq \sqrt{2cm}$ and Lemma <2>(v) bounds the final sum by a constant.

Proof (of Claim 1). In order to bound $|R^{(k)}(t)|_2$ we first need to focus on the behavior of the sequence $\{R_\alpha^{(k)}(t) : k \in \mathbb{N}\}$ for a fixed α. If $t_\alpha \leq 2^{-p[j,t]}$ then

$$R_\alpha^{(j)}(t) = R_\alpha^{(j-1)}(t) = \cdots = t_\alpha.$$

The first possible j for which $R_\alpha^{(j)}(t) < t_\alpha$ is the value for which

$$2^{-p[j,t]} < t_\alpha \leq 2^{-p[j-1,t]}.$$

Thereafter, we only know that $R^{(j)}_\alpha(t) \le 2^{-p[j,t]}$. Thus, by Lemma <2> parts (iii) and (iv),

$$
\begin{aligned}
|R^{(k)}(t)|_2^2 &= \sum\nolimits_{\alpha\in[\![M]\!]} \left(R^{(k)}_\alpha(t)\right)^2 \\
&\le \sum\nolimits_\alpha 2^{-2p[k,t]}\{t_\alpha > 2^{-p[k,t]}\} + \sum\nolimits_{j>k} t_\alpha^2\{2^{-p[j,t]} < t_\alpha \le 2^{-p[j-1,t]}\} \\
&\le 2^{k-2p[k,t]} + \sum\nolimits_{j\in\mathbb{N}}\{j>k\}2^{j-2p[j-1,t]}
\end{aligned}
$$

so that

\E@ Rk.norm <6>

$$|R^{(k)}(t)|_2 \le \sqrt{\sum\nolimits_{j\in\mathbb{N}}\{j\ge k\}2^{j-2p[j-1,t]}} \le \sum\nolimits_{j\in\mathbb{N}}\{j\ge k\}2^{j/2-p[j-1,t]}.$$

The final inequality comes from the general result: $\sqrt{\sum_j a_j} \le \sum_k \sqrt{a_j}$ for non-negative sequences $\{a_j\}$. □

Proof (of Claim 2). To bound the size of U_k, remember that $A^{(j)}(t)$ is non-zero only on the set $I_j(t) = \{\alpha \in [\![M]\!] : R^{(j-1)}(t) > 2^{-p[j,t]}\}$, a subset of the set $\mathcal{H}_j(t)$, which has size less than $R := 2^j$. There are at $2j$ values possible for $p[j,t]$ and for each α in $I_j(t)$ there are at most five values for $A^{(j)}_\alpha(t)$. Thus

\E@ Aj.vv <7>

$$\{A^{(j)}(t) : t\in T\} \subset \cup\{\mathcal{V}_{I,p} : |I|\le 2^j \text{ and } p\le 2j\}$$

where $\mathcal{V}_{I,p}$ denotes the set of all u in $\mathbb{R}^{[\![M]\!]}$ for which

$$2^p u_\alpha \begin{cases} \in\{0,1,2,3,4\} & \text{if } \alpha\in I \\ =0 & \text{if } \alpha\notin I \end{cases}.$$

For given I and p the set $\mathcal{V}_{I,p}$ has size $5^{|I|}$. Thus the union in <7> has size less than

$$2j\sum\nolimits_{\kappa\le 2^j}\binom{M}{\kappa}5^\kappa \le 2j(5M)^{2^j}\sum\nolimits_{\kappa\le 2^j}1/\kappa! \le M^{c_0 2^j} \qquad \text{if } M\ge 2$$

for some constant c_0. Consequently,

$$|U_k| \le \prod\nolimits_{j=1}^k |\{A^{(j)}(t) : t\in T\}| \le M^{c_0 2^1+\dots+c_0 2^k} \le M^{2c_0 2^k},$$

as asserted. □

References

Talagrand2021MMbook

Talagrand, M. (2021). *Upper and Lower Bounds for Stochastic Processes: Decomposition Theorems* (Second ed.), Volume 60 of *Ergebnisse der Mathematik und ihrer Grenzgebiete*. Springer-Verlag.