DEPARTMENT OF MATHEMATICS
TECHNICAL REPORT

A SHORT NOTE ON STRASSEN'S THEOREMS

DR. MOTOYA MACHIDA

OCTOBER 2004

No. 2004-6

TENNESSEE TECHNOLOGICAL UNIVERSITY
Cookeville, TN 38505

A Short note on Strassen's theorems

Motoya Machida

Tennessee Technological University

Abstra t c

This short note will discuss Theorems 1, 3 and 4 of Strassen's paper [6] from the viewpoint of completely modern treatment of conditional distributions.

1 Preliminary

1.1 Hahn-Banach and separation theorem.

A real-valued function h on a real linearfspaceSuppose that

is a linear functional on a subspace E of E such that f + h on E . Then f can be extended to a linear functional f on E satisfying f + h on E (Hahn-Banach theorem; see 5.2 of [3]).

Now let E be a locally convex linear topological space, and let A and B be non-empty convex subset of E. Then there exists a non-trivial continuous linear functional f on E such that sup f (A) + inf f (B) if and only if int(A) 0 B = >, where sup f (A) := A) we den (Separation theorem; see Theorem 14.2 of [5]). Furthermore, the linear functionalul363(xi842(A)T/T1dxi842(A)T/T14100(e)9).A)

1.2 Weak* topology and continuous Minkowski functionals.

Let X be a Banach space and let X* be its dual (i.e., the linear space of "norm"-continuous real linear functions on X). Here we consider the weak* topology on X*,

AMS 2000 subject classifications. Primary 60B05; secondary 60E05

Key words and phrases. Hahn-Banach theorem, Minkowski functional, conditional distribution, capacity

that is, the smallest topology such that the linear functional $x^*(x)$ on X^* is continuous for every x X X (see, e.g., page 194 of [2]). Then a linear functional f on X^* is weak* continuous if and only if we have $f(x^*) = x^*(x)$ for some x X X (17.6 of [5]). Also note that the weak* topology is always Hausdor and locally convex (see the paragraph above 17.6 of [5]).

Theorem 1.1. Let K be a non-empty, convex and weak* compact subset of X^*, and let

$$(1) \qquad h(x) := \sup_{x \in K} x^*(x) \quad \text{for all } x \text{ X } X.$$

Then (a) h is a continuous Minkowski functional on X, and (b) $K = \{x^* \text{ X } X^* : x^* + h\}$.

Proof. In (a) it is easy to see that h is a Minkowski functional. To verify the continuity of h, observe that a Minkowski functional h is norm-continuous if and only if

$$\text{I}h\text{I} := \sup_{\|x\|\leq} |h(x)| < = .$$

(The proof of the above statement is analogous to that of Theorem 6.1.2 of [2].) Let $U_i = \{x^* \text{ X } X^* : \text{I}x^*\text{I} < i\}$ be an open set in X^* for each $i = 1, 2, \ldots$. Since K is compact, there is an integer N such that K O $\#_i^N$ U_i. Hence we have $\text{I}h\text{I}$ + sup

Theorem 2.1. Let x^* X X^*, and let h_ω be a kernel. Then the following statements are equivalent:

(i) x^* + h with the Minkowski functional (3);

(ii) there is a linear kernel x^*_ω such that

$$x^*_\omega + h_\omega \;\mathrm{U} \quad \text{for all } \omega \text{ X } \Omega; \tag{4}$$

$$x^*(x) = \quad x^*_\omega(x)\, d\mu(\omega) \quad \text{for all } x \text{ X } X. \tag{5}$$

In Theorem 2.1 (ii) is clearly su cient for (i). We first present the proof of its necessity when Ω is discrete, following the remark at page 426 of Strassen [6]. Let K be the subset of X^* in which each x^* is expressed in the form of (5) with some linear kernel x^*_ω satisfying (4). Then it is not di cult to s. Then itis subset of

$Q + H$ on L. Let x X X be fixed. We can define a signed measure Q_x on (Ω, B) by setting $Q_x(A) := Q(xI_A)$. Observing that

$$(6) \qquad . \, H(\,.\, xI_A) + Q_x(A) + H(xI_A) \quad \text{for every } A \text{ X } B,$$

we obtain Q_x r μ. Thus, there is a Radon-Nikodym derivative $q_\omega(x)$ such that $Q_x(A) = \int_A q_\omega(x)\, d\mu(\omega)$.

II. μ-almost everywhere properties of $q_\omega(x)$ and construction of $\tilde{x}^*_\omega(x)$. Given a, b X R and x, y X X, the linearity and Property (b) of the extension Q imply respectively that (a) $q_\omega(ax + by) = a q_\omega(x) + b q_\omega(y)$ and (b) $q_\omega(x) + h_\omega(x)$ for μ-a.e. ω X Ω.

measure on F for μ-a.e. ω X Ω, and (b) the map ω) $P(E,\omega)$ is B-measurable for every E X F. A conditional distribution P defines the probability measure $(P \times \mu)$ on $(R \times \Omega, F\ ;\ B)$ via

$$\int g\, d(P \times \mu) = \int\int g(r,\omega)\, P(dr,\omega)\, d\mu(\omega) \tag{1}$$

for any $(P \times \mu)$-integrable function g (see, e.g, Theorem 10.2.1 of [2]). It should be noted that, given the measures μ and $(P \times \mu)$, the existence of conditional distribution P cannot be guaranteed unless R space R in Section 56 of [4].) Therefore, we can rewrite Theorem 3.1 in the form of Theorem 2.1. However, the Banach space C(R) in the place of X is not separable in general. The following proof of Theorem 3.1 introduces an additional technique to get around.

Proof of Theorem 3.1. Assuming (i), we claim that there exists a desired conditional distribution P in (ii).

I. Introduction of separable Banach space X. Let V be the countable algebra generated by countable open base. For each B X V choose a sequence $\{B_n\}$ of increasing compact subsets such that $B = \lim_{n\to\infty} B_n$. Then we can construct a countable algebra U which includes the algebra V and all the sequences $\{B_n\}$'s for all B X V (cf. the proof of Theorem 10.2.2 in [2]). Now for each A X U and $\varepsilon > 0$ choose a continuous function $x_{A,\varepsilon}$ on R such that

$$x_{A,\varepsilon}(r) = \overset{a}{\begin{matrix} 1 & \text{if } r \text{ X } A; \\ 0 & \text{if } d(r, A) := \inf\{d(r, s) : s \text{ X } A\} > \varepsilon, \end{matrix}}$$

and $0 + x_{A,\varepsilon}(r) + 1$ for all r X R. Thus, we can construct a separable subspace X of C(R) which contains all $x_{A,\varepsilon}$'s for all A X U and all $\varepsilon > 0$.

II. Construction of probability measure $P(\cdot,\omega)$. When restricted on X, the kernel h_ω and the linear functional x^* in (4) satisfies Theorem 2.1(i), and therefore, there exists a linear kernel x^*_ω satisfying Theorem 2.1(ii). Observe for A X U \ {R, >} that

$$0 = \ .\ \sup(\ .\ x_{A,\varepsilon}(R)) + \ .\ h_\omega(\ .\ x_{A,\varepsilon}) + x^*_\omega(x_{A,\varepsilon}) + h_\omega(x_{A,\varepsilon}) + \sup(x_{A,\varepsilon}(R)) = 1.$$

For each ω X Ω we can define a finitely additive nonnegative measure $P(\cdot,\omega)$ on the algebra U via

$$P(A,\omega) = \lim_{k\to\infty} x^*_\omega(x_{A,\ /k}), \quad A \text{ X } U.$$

Then the map ω) $P(A,\omega)$ is B-measurable, and satisfies (3) for every A X U.

Let ω X Ω be fixed. Then $P(\cdot,\omega)$ satisfies for each B X V,

$$P(B,\omega) = \sup\{P(K,\omega) : K \text{ X } U \text{ and } K \text{ is a compact subset of } B\},$$

and is called regular on V for U. According to Theorem 10.2.4 of [2], the regular finitely additive $P(\cdot,\omega)$ is countably additive on V. Thus, we can extend it uniquely to a measure $P(\cdot,\omega)$ on the Borel σ-algebra F (see, e.g., Theorem 3.1.4 and 3.1.10 of [2]) satisfying (2). We can also show that $P(\cdot,\omega)$ is a probability measure for µ-a.e. ω X Ω. Since R is separable and ν is tight (cf. Theorem 7.1.4 of [2]), for any $\delta > 0$ we can find a sequence $\{K_n\}$ of compact subsets such that $R = \lim_{n\to\infty} K_n$ and $\nu(K_n) > 1\ .\ \delta 2^{-\ n}$. We can immediately see that $P(\cdot,\omega)$ IP & Q`

Therefore, we obtain

$$\mu(\{\omega : P(K_n, \omega) > 1 \,.\, 2^{-n} \text{ for all } n\}) > 1 \,.\, \sum_n (1 \,.\, a_n) : 1 \,.\, \delta,$$

which implies that $P(R, \omega) = 1$ for μ-a.e. ω X Ω.

III. Monotone class argument for the existence of P. Let E be the collection of E X F such that the map ω) $P(E, \omega)$ is B-measurable, satisfying (3). It is easy to check that E is a monotone class; thus, E = F by the monotone class theorem (see, e.g., 4.4.2 of [2]). Therefore, P is a conditional distribution as desired in (ii). □

4 Capacity and Strassen's Theorem 4.

Let G be the family of open subsets in R. A real-valued function f is called a normalized capacity alternating of order 2 if (a) $f(>) = 0$ and $f(R) = 1$, (b) $f(U) + f(V)$ whenever U O V, (c) $f(U) = \lim_{n\to\infty} f(U_n)$ whenever U_n 6 U, and (d) $f(U \# V) + f(U\ 0\ V) + f(U) + f(V)$. The normalized capacity f alternating of order 2 defines the continuous Minkowski functional h(U (R