2. SUBADDITIVITY AND SUPERADDITIVITY

2.1. Geometric Subadditivity

Mathematics is filled with a range of tools which are at once simple and powerful. It is difficult to find a tool which surpasses subadditivity in terms of its combined simplicity and utility. Subadditive methods occupy a prominent position in this monograph and we begin by recalling the basic notions. Let $x_n,\ n \geq 1$, be a sequence of real numbers satisfying the "subadditive inequality"

$$x_{m+n} \leq x_m + x_n \quad \text{for all } m,\ n \in \mathbb{N}. \tag{2.1}$$

We say that the sequence $x_n,\ n \geq 1$, is *subadditive*. Subadditive sequences are nearly additive in the sense that they satisfy the *subadditive limit theorem*

$$\lim_{n \to \infty} \frac{x_n}{n} = \alpha$$

where $-\infty \leq \alpha < \infty$ and moreover

$$\alpha = \inf\{\frac{x_m}{m} : m \geq 1\}.$$

This is a standard result and we refer to Hille (1948) for a proof. The additive inverse of a subadditive sequence is a superadditive sequence, that is one which satisfies $x_{m+n} \geq x_m + x_n$ for all m and n. Subadditive sequences enjoy prominent use in various settings, including percolation and ergodic theory. They express the subadditivity of functions defined on the parameter set of intervals in $\mathbb{R}^+$.

Many sequences arising in applications are not subadditive but are "approximately subadditive" in the sense that they satisfy a generalized subadditive inequality of the form

$$x_{m+n} \leq x_m + x_n + \Delta_{m+n},$$

where $\Delta_k,\ k \geq 1$, is an appropriate sequence. If $\Delta_k,\ k \geq 1$, does not grow too fast then the sequence $x_n,\ n \geq 1$, still satisfies the subadditive limit theorem.

One of our central insights is that many graphs have an intrinsic subadditive and superadditive structure with respect to the parameter set of d-dimensional rectangles in $\mathbb{R}^d$. Although the subadditive structure expresses the self-similarity properties of the graph and is thus based on geometry in d dimensions, the analytic inequalities describing the corresponding graph length can usually be converted into inequalities involving approximately subadditive sequences. The intrinsic subadditivity of the traveling salesman graph, which will soon be made precise, was recognized by Beardwood, Halton, and Hammersley (1959). This observation formed the starting point for the proof of Theorem 1.1.

There are several ways to tease out the subadditive structure of graphs. For example, many functionals L^p occurring naturally in optimization problems satisfy a simple subadditivity condition

$$L^p(F \cup G) \leq L^p(F) + L^p(G) + C_1 t^p \tag{2.2}$$

for all finite sets F and G in $[0,t]^d$ where C_1 is a finite constant which may depend upon d and p. It is easy to check that the TSP, MST, and minimal matching functionals given by Definition 1.2 each satisfy simple subadditivity (2.2).

Simple subadditivity (2.2) is often stronger than what is needed. In most cases, it is sufficient to consider a modification of (2.2) which expresses an approximate subadditivity of L^p over the collection $\mathcal{R} := \mathcal{R}(d)$ of d-dimensional rectangles. If $R \in \mathcal{R}$ is partitioned into rectangles R_1 and R_2 then (2.2) gives for all finite sets F

$$L^p(F \cap R) \leq L^p(F \cap R_1) + L^p(F \cap R_2) + C_1(\text{diam}R)^p.$$

This condition expresses *geometric subadditivity* of L^p over rectangles. Unlike the subadditive relation (2.1), the subadditivity is only an approximate one and carries a correction term of $C_1(\text{diam}R)^p$.

Writing $L^p(F,R)$ for $L^p(F \cap R)$, we henceforth view L^p as a function defined on pairs (F,R) where F is a finite set and $R \in \mathcal{R}$ is a d-dimensional rectangle. This notation is not accidental and in fact explicitly recognizes that L^p is a function of two arguments, a point of view which is central to the development of our subject. With this notation we obtain

$$L^p(F,R) \leq L^p(F,R_1) + L^p(F,R_2) + C_1(\text{diam}R)^p. \tag{2.3}$$

Although we emphasize that (2.3) represents only approximate subadditivity, we will sometimes omit mention of "approximate" and simply refer to (2.3) as geometric subadditivity. The relation (2.3) may appear weak but we will see that when F represents a random set it can be turned into an inequality involving sequences which are approximately subadditive. This device will form the foundation of our approach.

By iterating (2.3) it is straightforward to check that if $\{Q_i\}_{i=1}^{2^{dj}}$ is a partition of $[0,1]^d$ into subcubes of edge length 2^{-j} then for $0 < p < d$ we have

$$L^p(F,[0,1]^d) \leq \sum_{i=1}^{2^{dj}} L^p(F \cap Q_i, Q_i) + C_1' 2^{(d-p)j}, \tag{2.4}$$

where the constant C_1' depends now on d and p. Indeed, when $j = 1$ we obtain for the partition $\{Q_i\}_{i=1}^{2^d}$

$$L^p(F,[0,1]^d) \leq \sum_{i=1}^{2^d} L^p(F \cap Q_i, Q_i) + C_1 2^d D^p,$$

where $D := \text{diam}[0,1]^d = d^{1/2}$. When $j = 2$ we obtain

$$\begin{aligned} L^p(F,[0,1]^d) &\leq \sum_{i=1}^{2^{2d}} L^p(F \cap Q_i, Q_i) + \sum_{i=1}^{2^d} 2^d \cdot C_1 \cdot (D/2)^p + \\ &\quad + C_1 2^d D^p \\ &\leq \sum_{i=1}^{2^{2d}} L^p(F \cap Q_i, Q_i) \ + \ C_1 2^d \cdot D^p(2^{d-p}+1). \end{aligned}$$

Iterating j times gives

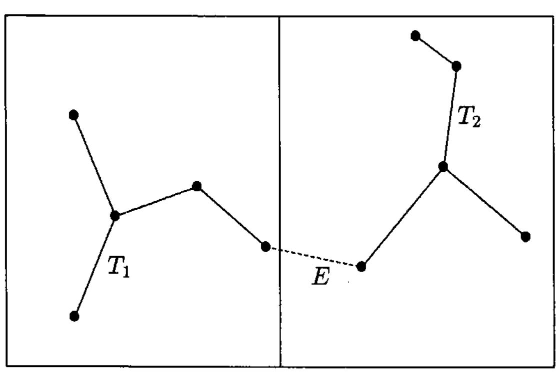
$$L^p(F,[0,1]^d) \le \sum_{i=1}^{2^{jd}} L^p(F \cap Q_i, Q_i) + \\ + \; C_1 2^d \cdot D^p(2^{(j-1)(d-p)} + ... + 2^{d-p} + 1)$$

which is exactly (2.4).

The subadditive relation (2.3) is simpler than the customary one appearing in problems of this sort. One usually requires that (2.4) hold whenever Q_i, $i \ge 1$, are congruent subcubes with edge length $m^{-1/d}$, so that the correction term on the right side of (2.4) is $C_1 m^{(d-p)/d}$. Checking such a condition is a little more involved than checking the relation (2.3). We have chosen the form (2.3) of subadditivity for three reasons: the first is that it will conveniently accommodate problems defined on domains larger than the unit cube, the second is that it will parallel an analogous definition of superadditivity, and the third is its evident simplicity.

Many of the classic optimization problems satisfy geometric subadditivity (2.3). For example, to see that the length of the minimal spanning tree M^p is subadditive (2.3) we argue as follows: given a finite set F and a rectangle $R := R_1 \cup R_2$, let T_i denote the minimal spanning tree which realizes $M^p(F \cap R_i, R_i)$, $1 \le i \le 2$. Tie together the local spanning trees T_1 and T_2 with an edge which has a length bounded by the sum of the diameters of the rectangles R_1 and R_2 (see Figure 2.1). Performing this operation generates a feasible tree on F at a total cost bounded by the right side of (2.3). Now (2.3) follows by minimality. Showing that the TSP and minimal matching problem satisfy (2.3) involves similar considerations.

Figure 2.1. The MST is subadditive: the length of the global MST is bounded by the lengths of the local trees T_1 and T_2 and a connecting edge E



2.2. Geometric Superadditivity, Boundary Functionals

Subadditive relations (2.3) are significantly strengthened when coupled with superadditive relations. There are several reasons to search for superadditivity. The most compelling one is that if we can show superadditivity together with the approximate subadditivity (2.4), then the functional L^p becomes "nearly additive" in the sense that

$$L^p(F \cap R) \approx L^p(F \cap R_1) + L^p(F \cap R_2).$$

Relations of this sort are crucial in showing that a global graph length can be approximately expressed as a sum of the lengths of local components.

Unfortunately, most optimization functionals L^p lack an intrinsic superadditive property which leads to interesting limit results. This drawback motivated Steele (1981a) to define a condition which can occasionally serve as a substitute for superadditivity. His condition, termed "upper linearity", can be put to use in some situations.

The most convenient and elegant way to circumvent the lack of superadditivity involves an appropriate modification of the functionals L^p. The behavior of functionals on point sets in the unit cube is almost always influenced by the boundary of the unit cube. This presents annoying technical and conceptual difficulties and can sometimes be overcome by identifying the opposite faces of the cube. This amounts to replacing the usual metric on the cube by the "flat metric" and is usually an undesirable oversimplification.

We use this apparently annoying boundary effect to our advantage. We will use the boundary effects to modify the original problem into a "boundary problem". Boundary problems give rise to "boundary graphs", whose lengths we refer to as "boundary functionals". "Boundary functionals" will have precisely the sought after superadditivity. Roughly speaking, boundary functionals, denoted here by L_B, are defined on pairs $(F, R) \in \mathcal{F} \times \mathcal{R}$, and measure the length of those graphs on the vertex set F together with the boundary of R, which is treated as a single point. Thus edges on the boundary of R have zero length. In the case of the traveling salesman functional, this means that travel on the boundary ∂R is free. Implicit in their definition is the fact that boundary functionals L_B are smaller than the standard functional L.

The boundary functional $L_B(F, R)$ treats the boundary of R as a single point so that all edges joined to the boundary are joined to one another. Cognoscenti will recognize that boundary functionals are analogous to the "wired boundary condition" used in percolation and statistical mechanics. They are also analogous to the "wired spanning forest" used in the study of random trees (cf. Lyons and Peres, 1997).

As we will see shortly, boundary functionals L_B, which represent a slight modification of the underlying optimization functional L, are intrinsically superadditive. L is intrinsically subadditive. It follows that L is "nearly additive" whenever L_B is a "close" approximation to L.

This crucial feature makes boundary functionals a natural choice of study and, as we will see in the sequel, leads to a wealth of asymptotic estimates. In many instances, asymptotics for optimization functionals can be deduced from the study of asymptotics for boundary functionals. This idea is at the heart of our subject and will appear over and over again.

If $R \in \mathcal{R}$ is partitioned into rectangles R_1 and R_2 then a typical boundary functional L_B satisfies

$$L_B(F,R) \ge L_B(F \cap R_1, R_1) + L_B(F \cap R_2, R_2) \tag{2.5}$$

and thus L_B is *superadditive with no error term.* The absence of an error term in (2.5) stands in sharp contrast to the approximate subadditive relation (2.3). In the sequel we will convert (2.5) into an inequality involving sequences which are superadditive and not merely "approximately superadditive". This distinction has telling consequences and will play a central role in our development of limit theorems for optimization functionals. This will be developed more fully in Chapter 4.

We illustrate the idea of boundary functionals by formally describing some concrete examples. Other examples will follow in the sequel.

The Boundary TSP Functional

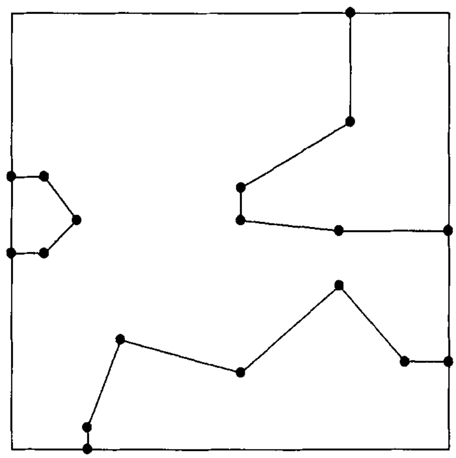
For all d-dimensional rectangles $R \in \mathcal{R}$, finite sets $F \subset R$, and $p \ge 1$, let $T^p(F,R,(a,b))$ denote the length of the shortest path with pth power weighted edges through $F \cup \{a,b\}$ with endpoints a and b, where a and b belong to ∂R. Define the boundary functional T_B^p associated with T^p by

$$T_B^p(F,R) := \min\left(T^p(F,R), \quad \inf \sum_i T^p(F_i,R,(a_i,b_i))\right),$$

where the infimum ranges over all partitions $(F_i)_{i \ge 1}$ of F and all sequences of pairs of points $(a_i,b_i)_{i \ge 1}$ belonging to ∂R.

The boundary functional $T_B^p(F,R)$ may be interpreted as the length of an optimal cycle (with pth power weighted edges) through the set F which may repeatedly exit to the boundary of R at one point and re-enter at another, incurring no cost when moving along the boundary. See Figure 2.2. It is easy to check that T_B^p satisfies superadditivity (2.5): for each $1 \le i \le 2$, the restriction of the global tour $T_B^p(F,R)$ to the rectangle R_i defines a boundary tour of the set $F \cap R_i$, which by minimality is at least as large as $T_B^p(F,R_i)$. This is precisely (2.5). Superadditivity breaks down for the functional T_B^p when $0 < p < 1$.

Figure 2.2. The boundary TSP graph



It is likewise easy to show that T_B^p satisfies simple subadditivity (2.2). The boundary TSP functional T_B^p is thus a modification of the standard TSP functional T^p. The boundary functional would not be so interesting were it not for its close approximation of the standard functional T^p. This closeness property is examined further in the next chapter.

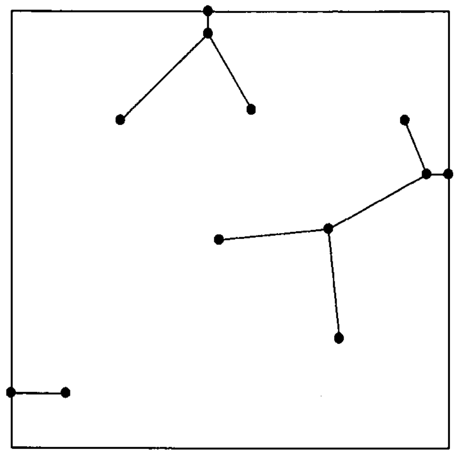
The Boundary MST Functional

For all d-dimensional rectangles $R \in \mathcal{R}$, finite sets $F \subset R$, and $p \geq 1$, define the boundary MST functional by

$$M_B^p(F,R) := \min\left(M^p(F,R), \quad \inf \sum_i M^p(F_i \cup a_i)\right),$$

where the infimum ranges over all partitions $(F_i)_{i \geq 1}$ of F and all sequences of points $(a_i)_{i \geq 1}$ belonging to ∂R. When $M_B^p(F,R) \neq M^p(F,R)$ the graph realizing the boundary functional $M_B^p(F,R)$ may be thought of as a collection of small trees connected via the boundary ∂R into a single large tree, where the connections on ∂R incur no cost. See Figure 2.3. It is a simple matter to see that the boundary MST functional is simply subadditive (2.2) and also superadditive (2.5). Later we will see that the boundary MST functional closely approximates the standard MST functional M^p.

Figure 2.3. The boundary MST graph



The Boundary Minimal Matching Functional

For all d-dimensional rectangles $R \in \mathcal{R}$, finite sets $F \subset R$, and $p \geq 1$, we let $S_B^p(F, R)$ denote the length of the least Euclidean matching (with pth power weighted edges) of points in F with matching to points on ∂R permitted. More precisely, each point in F is paired with either a boundary point on ∂R or another point in F; $S_B^p(F, R)$ minimizes the sum of the pth powers of the edge lengths over all such pairings. We allow the possibility that even when F has even parity, one point in F may be isolated and unmatched. This is needed to ensure superadditivity and to account for the case that minimal matching on an odd number of points leaves one point unmatched. See Figures 2.4 and 2.5.

As is the case with the boundary TSP and MST functionals, it is easy to verify that the boundary minimal matching functional is simply subadditive (2.2) and superadditive (2.5). See Figure 2.5. Later we will see that the boundary minimal matching functional closely approximates the standard minimal matching functional S^p.

Figure 2.4. The boundary minimal matching graph

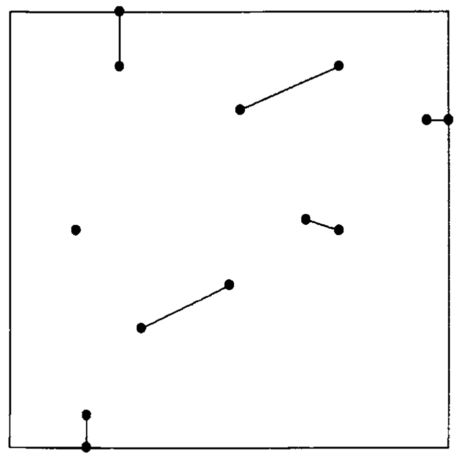
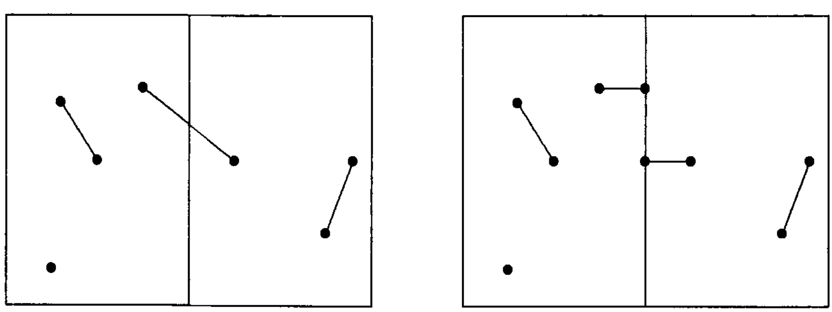


Figure 2.5. Boundary minimal matching is superadditive: the length of the global matching (left) exceeds the length of the two local matchings (right)



In the above examples, the boundary functional measures the minimal length of a graph satisfying a specified boundary condition. We have restricted attention to the range $p \geq 1$ since without further modifications the functionals would no longer enjoy superadditivity.

Notes and Notation

(i) If L is an optimization functional, then L_B will henceforth designate the canonical boundary functional. We will often abbreviate the notation and write $L^p(F)$ for $L^p(F,[0,1]^d)$ and similarly for $L^p_B(F)$. When $p = 1$ we will suppress mention of p and simply write $L(F,R)$ and $L_B(F,R)$ in place of $L^1(F,R)$ and $L^1_B(F,R)$, respectively. Notice that $L^p(F,\mathbb{R}^d) = L^p_B(F,\mathbb{R}^d)$.

Many of the subadditive functionals L^p considered in this monograph satisfy $L^p(F,R_1) = L^p(F,R_2)$ if $F \cap R_1 = F \cap R_2$. On the other hand, the superadditive functionals exhibit a higher dependence on the underlying rectangular domain and in general will *not* satisfy $L^p(F,R_1) = L^p(F,R_2)$ when $F \cap R_1 = F \cap R_2$.

(ii) We use L_B to designate the canonical boundary functional. Previous work of Redmond and Yukich (1994, 1996) referred to boundary functionals as "rooted duals" and denoted them by L_r.