CHAPTER 4

Multiple dimensions: Weighted Euclidean distance

The whole of science is nothing more than a refinement of everyday thinking.

(Albert Einstein, *Physics and Reality,* 1936)

In Chapter 3, we outlined the logic of complex policy spaces in terms of "everyday thinking." In this chapter, we will refine that material, using the language of matrix notation. Instead of two dimensions, however, we will assume that the policy space has n dimensions, where n is an arbitrary number. Technically, the policy space $\mathcal{P}$ is the generalized Cartesian product: $\mathcal{P} = P_1 \times P_2 \times \cdots \times P_n$.

As before, each dimension measures the budget of one project, and members are assumed to have preferences over all the projects. Using matrix notation will allow us to make precise the ideas of *salience* and *nonseparability* that we presented graphically in Chapter 3. Mathematical notation will afford us several refinements that make the additional complexity worthwhile. Matrices and vectors are useful notational devices because they save space and clutter in representing organized arrays of characters. Once you get used to the idea of using matrix notation, you will find it *easier* than using summations cluttered with indexes and brackets.

Matrix notation and definitions

- A *matrix* is an ordered array of numbers or characters. We will use **boldface** type to distinguish a matrix from *scalar* variables. A scalar might be thought of as a 1×1 matrix, but there is no need to be so fussy. Any real matrix has size $r \times c$, or number of rows times num-

Note: This chapter contains technical details of the discussion in Chapter 3 and can be skipped without loss of continuity. However, the reader interested in technical details is advised to study this chapter carefully.

ber of columns. For example, a 4×6 matrix **A** will have twenty-four entries, arranged in four rows and six columns.

$$\mathbf{A} = \begin{array}{cc} \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ a_{31} & a_{32} & a_{33} & a_{34} & \boxed{a_{35}} & a_{36} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} \end{bmatrix} & \begin{array}{c} \uparrow \\ \text{4 rows} \\ \downarrow \end{array} \\ \leftarrow \text{ 6 columns } \rightarrow & \end{array} \tag{4.1}$$

Specific entries are denoted by subscripts, with the first subscript representing the row, and the second subscript representing the column. For example, a_{35} is the element of **A** in the *third* row and *fifth* column (see box in figure above).

- There is one special kind of matrix, called a *vector,* that we will find especially useful. A vector may have only one row (a *row* vector) or only one column (a *column* vector). That is, row vectors are $1 \times c$, column vectors are $r \times 1$. Consider these examples of a row vector **R** and a column vector **C**.

$$\mathbf{R} = [a_{11}\ a_{12}\ a_{13} \cdots a_{1c}] \quad \mathbf{C} = \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \\ \vdots \\ a_{r1} \end{bmatrix} \tag{4.2}$$

- A *diagonal* matrix is all zeros, except for the elements along the main diagonal (more precisely, $a_{jk} = 0$ for all $j \neq k$). A special case of a diagonal matrix is the *identity* matrix, which consists of 1s along the main diagonal and zeros everywhere else. Here is an example of a diagonal matrix **D** and an identity matrix **I**:

$$\mathbf{D} = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 9 \end{bmatrix} \quad \mathbf{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4.3}$$

Basic matrix operations

- *Transpose:* The transpose of a matrix is another matrix, with rows becoming columns and vice versa. The matrix being transposed has a superscript T. Transposition is easy in the case of row and column vectors: If the number of rows in **R** is equal to the number of columns in **C**, and they have the same elements, then $\mathbf{R}^T = \mathbf{C}$, and $\mathbf{C}^T = \mathbf{R}$. More generally, suppose an arbitrary matrix **M** is size $r \times c$. Then $\mathbf{M}^T$ will be size $c \times r$.

For example, consider the following matrix **M** and its transpose $\mathbf{M}^T$:

$$\mathbf{M} = \begin{bmatrix} 6 & 9 \\ -1 & 4 \end{bmatrix} \quad \mathbf{M}^T = \begin{bmatrix} 6 & -1 \\ 9 & 4 \end{bmatrix} \tag{4.4}$$

Notice that the first *column* [6–1] has become the first *row*, and so on. Importantly, this means that $\mathbf{D}^T = \mathbf{D}$ if and only if **D** is diagonal.

- *Addition:* Matrices can be added only if they are the same size. We can identify the sum of **A** and **B** as the new matrix **C**, so that $\mathbf{A} + \mathbf{B} = \mathbf{C}$. The elements of **C** are the sums of the respective elements of **A** and **B**, as the following 2×2 example shows:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} (a_{11} + b_{11}) & (a_{12} + b_{12}) \\ (a_{21} + b_{21}) & (a_{22} + b_{22}) \end{bmatrix} \tag{4.5}$$

Subtraction is defined the same way, except that the elements of the resulting matrix are the *difference* of the elements of the matrices being subtracted.

- *Multiplication:* Multiplying a scalar times a matrix is easy and works just the way one would expect. Suppose you have a scalar s and a matrix **A**; then their product would be:

$$s\mathbf{A} = s \times \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} s \times a_{11} & s \times a_{12} \\ s \times a_{22} & s \times a_{22} \end{bmatrix} \tag{4.6}$$

Multiplying two matrices is a little harder. It is useful to keep three things in mind for matrix multiplication. First, make sure the number of *columns* of the matrix on the left is the same as the number of *rows* of the matrix on the right. If this condition is not met, multiplication is not defined for these two matrices. The reason that this condition is important is that each element in the resulting matrix is the sum of the products of the elements across the columns of the matrix on the left and the elements down the rows of the matrix on the right. Second, the size of the matrix representing the product of any two matrices will be the number of *rows* in the first matrix and the number of *columns* in the second matrix. Obviously, matrix multiplication is very different from scalar multiplication, because the order in which the matrices are multiplied matters. Third, matrix multiplication is associative, just as in scalar multiplication, for a given order of matrices to be multiplied. For example, $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$, just as in scalar multiplication.

Consider the following three examples:

Multiplication Example 1: Two vectors make a scalar (1×3 times $3 \times 1 = 1 \times 1$, or a scalar).

$$\begin{aligned} [5 \;\; 1 \;\; 3]\begin{bmatrix} 6 \\ -2 \\ 2 \end{bmatrix} &= (5 \times 6) + (1 \times -2) + (3 \times 2) \\ &= 34 \end{aligned} \tag{4.7}$$

Multiplication Example 2: The same two vectors make a matrix, if multiplied in the opposite order (3×1 times $1 \times 3 = 3 \times 3$ matrix).

$$\begin{bmatrix} 6 \\ -2 \\ 2 \end{bmatrix} [5 \;\; 1 \;\; 3] = \begin{bmatrix} 30 & 6 & 18 \\ -10 & -2 & -6 \\ 10 & 2 & 6 \end{bmatrix} \tag{4.8}$$

Multiplication Example 3: Two matrices make another matrix, with the elements of the product being the sum of the products of the elements across the columns of the first matrix and down the rows of the second matrix. Specifically, for the two matrices multiplied below, the top left element (a_{11}) in the product matrix is $(1 \times 2) + (3 \times 8) + (5 \times 14) = 96$, and so on.

$$\begin{bmatrix} 1 & 3 & 5 \\ 7 & 9 & 11 \\ 13 & 15 & 17 \end{bmatrix} \times \begin{bmatrix} 2 & 4 & 6 \\ 8 & 10 & 12 \\ 14 & 16 & 18 \end{bmatrix} = \begin{bmatrix} 96 & 114 & 132 \\ 240 & 294 & 348 \\ 384 & 474 & 564 \end{bmatrix} \tag{4.9}$$

We are now ready to apply the concepts of matrices and vectors in refining the intuitive exposition of Chapter 3.

SED and WED: Spatial preferences in two-dimensional space

As the reader will recall from Chapter 2, the problem of "representing" preferences for citizens is a very general one. In this book, we have settled on "spatial" representations of preferences, meaning that the distance between a voter's ideal point and a proposal is inversely related to utility. More simply, voters like a proposal less as it differs more from what they want.

But now we must define "distance," or the difference between proposals, in several dimensions rather than just one. Generally, the simplest measure of distance is *simple Euclidean distance* (SED). In a single dimension, the SED between two points y and z is just the absolute value of the difference. Consider *representative citizen i*, who has an *ideal point* x_i. In Chapter 2, we used the absolute value of the difference between budgets to define preference (*i prefers* y to z if and only if $|y - x_i| < |z - x_i|$) and indifference (*i* is *indifferent* between y and z if and only if $|y - x_i| = |z - x_i|$).

The SED between two points in two-dimensional space is more complicated to compute, but has the same meaning. We can define the points in two-dimensional space, using vectors. An element in our two-dimensional policy space will be represented as a column vector with the Project 1 budget on top and the Project 2 budget below it. For example,

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \qquad \mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \tag{4.10}$$

are two alternatives (**y** and **z**) in the policy space.

Each member's ideal point is also a vector, with an ideal budget for each project. How are we to represent the distances from ideal points to budgets using vectors? The difference between the vectors is just another vector, giving the difference in the budget *in each project.* But a vector is not, itself, a distance. To compute distance, we apply the Pythagorean theorem. Imagine we are given an arbitrary right triangle with sides measuring α and β and hypotenuse measuring γ. The Pythagorean theorem says that $\gamma^2 = \alpha^2 + \beta^2$. Solving for γ, we get:

$$\gamma = \sqrt{\alpha^2 + \beta^2} \tag{4.11}$$

As Figure 4.1 shows, the "difference" between proposals **y** and **z** has two parts: The difference between the two projects has two dimensions. For Project 1, the difference is measured on the horizontal axis (distance α). For Project 2, the difference is measured on the vertical axis (distance β). The *distance* between the two proposals, however, is the hypotenuse γ (dotted line) of the triangle. Consequently, the formula

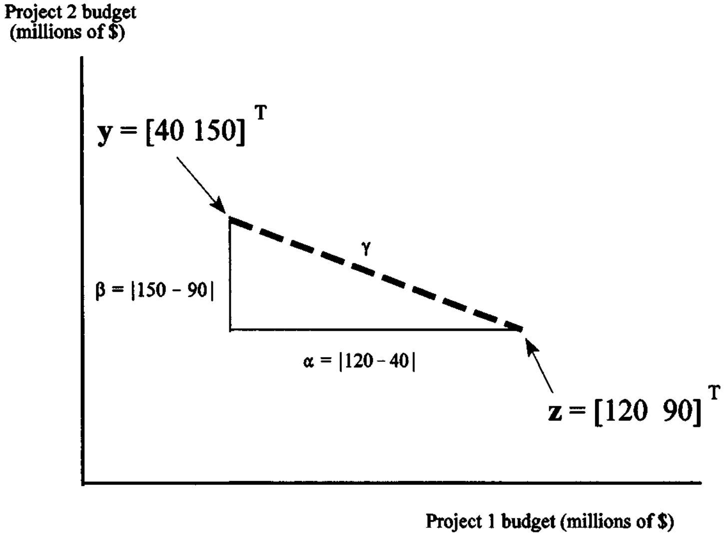

Figure 4.1. Computation of simple Euclidean distance in two-dimensional space.

for SED(**y**, **z**) in two-dimensional space is derived directly from the Pythagorean theorem.

$$\mathrm{SED}(\mathbf{y}, \mathbf{z}) = \sqrt{(y_1 - z_1)^2 + (y_2 - z_2)^2} \tag{4.12}$$

More generally, for an arbitrary number of dimensions n, the distance between two points **y** and **z** is:

$$\mathrm{SED}(\mathbf{y}, \mathbf{z}) = \sqrt{\sum_{j=1}^{n}(y_j - z_j)^2} \tag{4.13}$$

That is, SED is the square root of the sum of the squared differences in each respective dimension. Using matrix notation, we can rewrite (4.13) as:

$$\mathrm{SED}(\mathbf{y}, \mathbf{z}) = \sqrt{[\mathbf{y} - \mathbf{z}]^{\mathrm{T}}[\mathbf{y} - \mathbf{z}]} \tag{4.14}$$

Here is an illustration of the advantage of matrix notation: Equation (4.13) is different for any number of dimensions, since n (the number of issues) differs. But n does not appear in (4.14); matrix notation is more general and applies to any number of dimensions without modification.

SED is the simplest way to represent preferences in spatial theory, but not the only way. A trivial generalization of SED is the set of order-preserving "affine" transformations of SED. Order-preserving affine transformations are linear rescalings that keep the same direction of increase from small to large. We say y is an affine transformation of x if

$$y = h + (k \times x) \tag{4.15}$$

where h and k are arbitrary scalars ($k > 0$). If $x_1 > x_2$, then a transformation of x is order preserving if $y_1 = h + (k \times x_1)$ and $y_2 = h + (k \times x_2)$ have the same order: $y_1 > y_2$. SED or any affine transformation of SED is equally valid as a representation of preference. To put it more simply, requiring that preferences are spatial does not restrict us to any one preference function.

This "nonuniqueness" seems strange, but it is not an uncommon situation in measurement problems. For example, the concept of temperature can be conceived as a large ordered set, whose elements range from very cold to very hot. The specific units we use to "measure" temperature, as well as the zero point in the scale, are generally arbitrary, however.[1] Zero degrees Fahrenheit is quite arbitrary as a level of temperature. Zero degrees Celsius corresponds to 32° Fahrenheit; 100° Celsius is the same as 212° Fahrenheit. Consequently, both Celsius and Fahrenheit measure the same phenomenon (ambient heat energy), so clearly neither is unique. Measurement theorists say the two scales "are unique only up to an affine transformation." Simple algebra quickly shows the form of the affine transformation – one way of converting between Celsius and Fahrenheit is:

$$C° = 5/9\,(F° - 32) \approx -\,17.7778 + 0.556F° \tag{4.16}$$

Using the general definition of affine transformations in (4.15), we see that in the Fahrenheit–Celsius example $h = -17.78$, and $k = 0.556$.

We want to represent another very complicated concept: political preferences. If we use SED, we have to remember that (as with temperature) the units are arbitrary and need not be "anchored" at any fixed zero point. Differences in satisfaction are only important ordinally. Or-

 https://doi.org/10.1017/CBO9781139174725.006

der-preserving affine transformations don't change preferences in any interesting way: SED or $65 + (4.1698 \times \text{SED})$ are equally valid measures of distance between alternatives as a representation of preferences.

There is another generalization of SED called "weighted Euclidean distance" (WED). To understand why WED is a genuine generalization, rather than just a transformation, notice that SED (and any affine transformation of SED) requires that preferences have two characteristics: (1) *Preferences must be separable* (the expected *level* of one issue doesn't affect the conditional *ideal* for other issues). (2) *All dimensions have equal salience* (distances in all issues "count" the same). Unlike a simple affine transformation of SED, WED relaxes both these requirements. Consequently, WED is substantively more general than SED.

The formula for WED, assuming that the matrix of weights $\mathbf{A}$ is voter specific (i.e., subscripted by i), is:

$$\text{WED}(\mathbf{y}, \mathbf{z}) = \sqrt{[\mathbf{y} - \mathbf{z}]^{\mathrm{T}} \mathbf{A}_i [\mathbf{y} - \mathbf{z}]} \tag{4.17}$$

Compare Equation (4.17) to Equation (4.14): The difference between SED and WED is the matrix of salience and interaction terms $\mathbf{A}_i$. If there are two projects, $\mathbf{A}_i$ would look like this:

$$\mathbf{A}_i = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \tag{4.18}$$

We can distinguish two types of terms, salience and interaction terms, in the matrix of preference weights $\mathbf{A}_i$.

- The *main diagonal* elements, a_{11} and a_{22}, are the salience terms. Salience is a measure of the relative value the voter places on Project 1 and Project 2, given the scale of those projects (e.g., dollars or number of people). Salience terms are assumed to be nonnegative by definition, and in practical terms we will consider only positive salience terms. (A salience term equal to zero means the voter doesn't ever consider this project or policy in making decisions.)
- The *off-diagonal* elements are called interaction terms. The interactions measure how much the voter's evaluation of *changes* in one project depend on the expected *level* of another project. For convenience we will generally assume that $a_{12} = a_{21}$, though there is nothing in the theory itself that requires this restriction.[2]

If $\mathbf{A}_i = \mathbf{I}$ (meaning that $a_{11} = a_{22} = 1$, and $a_{12} = a_{21} = 0$), then SED(**y**, **z**) = WED(**y**, **z**). It is useful to define symmetric preferences for the multidimensional setting:

Symmetric preference. Preferences of voter i are symmetric if and only if the voter likes equally well any two points **y** and **z** that (a) lie on a straight line passing through $\mathbf{x}_i$ and (b) are equidistant from $\mathbf{x}_i$ (in terms of SED).

Like SED preferences, WED preferences are symmetric.

We can also define the concept of indifference more generally.

Indifference. A voter i is indifferent between any two alternatives **y** and **y**′ such that $\text{WED}(\mathbf{y} - \mathbf{x}_i) = \text{WED}(\mathbf{y}' - \mathbf{x}_i)$. In general, we can identify a set of alternatives **y** among which the voter is indifferent: $\{\mathbf{y} \mid \text{WED}(\mathbf{y} - \mathbf{x}_i) = K\}$. The variable K is a constant whose value can range from zero (at the ideal point) to an arbitrarily large number. The graph of the elements of this set of points among which the voter is indifferent is called an "indifference curve." Each distinct value of K defines one indifference curve for the voter. Since any particular K defines one indifference curve, it follows that each point in the policy space $\mathcal{P}$ is located on exactly one indifference curve. Consequently, there can be no intersection of indifference curves in preferences measured by WED, just as was true for SED.

Consider the following possibilities for different kinds of indifference curves.

- If $\mathbf{A}_i = k\mathbf{I}$ (k an arbitrary scalar, $\mathbf{I}$ an identity matrix of order n), then indifference curves are *circles* with centers at the voter's ideal point, because issues are weighted equally and preferences are separable.
- For any $\mathbf{A}_i = \mathbf{sI}$ (where $\mathbf{s}$ is a $1 \times n$ row vector of salience weights, so that $\mathbf{A}_i$ is diagonal), indifference curves are *ellipses* centered at the ideal point, with major axes parallel to the policy dimensions. The reason is that the voter values issues differently, so equal departures from his ideal have different impacts on his utility, depending on the project.
- For any $\mathbf{A}_i$ that is not diagonal, preferences are not separable. As we

 https://doi.org/10.1017/CBO9781139174725.006

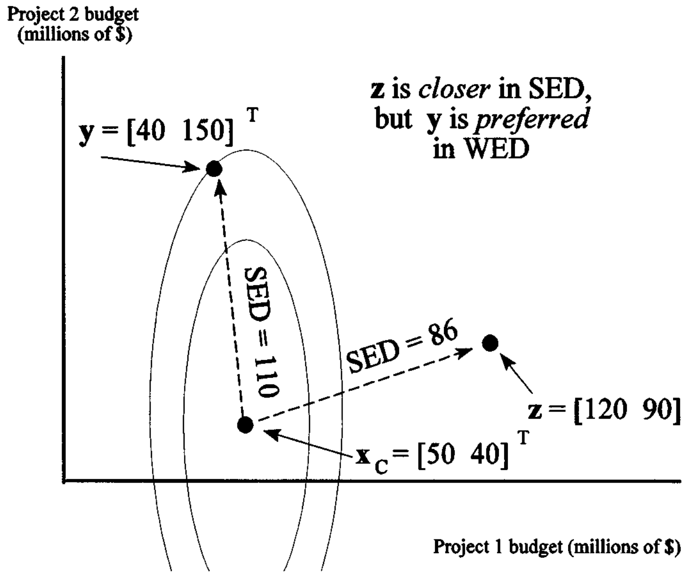

Figure 4.2. Comparison of **y** and **z** using WED for Member C.

saw in Chapter 3, this means that the form of the preference ellipses is complex, with the major axis of the ellipse tilted at an angle rather than parallel to one of the policy dimensions.

An example using (separable) WED in two-dimensional space

Returning to the preferences of the subcommittee members, consider first C's preference rule. We will say he prefers **y** to **z** if and only if **y** is closer to his ideal point, or $\text{WED}_A(\mathbf{y}, \mathbf{x}_C) < \text{WED}_A(\mathbf{z}, \mathbf{x}_C)$, where as before $\mathbf{y} = [40\ 150]^T$, $\mathbf{z} = [120\ 90]^T$, and $\mathbf{x}_C = [50\ 40]^T$. As Figure 4.2 shows, Mr. C is indifferent among alternatives that lie exactly K units (in WED, not SED) from his ideal point.

Assume that the $\mathbf{A}_C$ matrix looks like this:

$$\mathbf{A}_C = \begin{bmatrix} 15 & 0 \\ 0 & 1 \end{bmatrix} \tag{4.19}$$

Mr. C's preferences for Projects 1 and 2 are separable (you can tell because $a_{12} = a_{21} = 0$). The issue salience values are very different, however: Each dollar difference from his ideal in Project 1 "counts" more because $a_{11} > a_{22}$.

Given the weights in $\mathbf{A}_C$, does C prefer $\mathbf{y}^T = [40, 150]$ or $\mathbf{z}^T = [120, 90]$, remembering $\mathbf{x}_C^T = [50, 40]$? For comparison, first calculate SED($\mathbf{y}$, $\mathbf{x}_C$) and SED($\mathbf{z}$, $\mathbf{x}_C$):

$$\begin{aligned}\mathrm{SED}(\mathbf{y}, \mathbf{x}_C) &= \sqrt{[-10\ 110]\begin{bmatrix}-10\\110\end{bmatrix}}\\ &= \sqrt{(-10)^2 + 110^2} = 110.45\end{aligned} \tag{4.20}$$

$$\mathrm{SED}(\mathbf{z}, \mathbf{x}_C) = \sqrt{[70\ 50]\begin{bmatrix}70\\50\end{bmatrix}} = \sqrt{70^2 + 50^2} = 86.02 \tag{4.21}$$

The definition of WED implies that if $\mathbf{A}_C$ were an identity matrix, WED and SED are identical. Equations (4.20) and (4.21) would then establish that C prefers $\mathbf{z}$ to $\mathbf{y}$. For $\mathbf{A}_C \neq \mathbf{I}$ (as given in Equation (4.17)), however, we must use WED:

$$\begin{aligned}\mathrm{WED}(\mathbf{y}, \mathbf{x}_C) &= \sqrt{[-10\ 110]\begin{bmatrix}15 & 0\\0 & 1\end{bmatrix}\begin{bmatrix}-10\\110\end{bmatrix}}\\ &= \sqrt{(-10^2) \times 15 + 110^2} = 116.62\end{aligned} \tag{4.22}$$

$$\begin{aligned}\mathrm{WED}(\mathbf{z}, \mathbf{x}_C) &= \sqrt{[70\ 50]\begin{bmatrix}15 & 0\\0 & 1\end{bmatrix}\begin{bmatrix}70\\50\end{bmatrix}}\\ &= \sqrt{(70^2) \times 15 + 50^2} = 275.68\end{aligned} \tag{4.23}$$

Now, because of the large weight attached to Project 1, C prefers proposal $\mathbf{y}$, though $\mathbf{z}$ appears closer to $\mathbf{x}_C$ in terms of simple distance. The reason can be recalled from Figure 4.2: Member C cares more about small differences in the Project 1 budget. Consequently, horizontal movements cause a larger decline in utility as distance from $\mathbf{x}_C$ increases.

What if preferences are nonseparable?

Allowing preferences to be nonseparable makes the process of modeling vote choice much more complex, but also more realistic. Recall the "paradox" from the preceding chapter:

The paradox of nonseparability. Suppose a voter has an unconditional "package" ideal point $\mathbf{x}_i$, and that preferences are nonseparable. Now, imagine that the position of the committee on issue 1 is fixed at $\tilde{x}_1$. The voter may actually prefer, and vote for, movements *away from* $\mathbf{x}_i$ on votes that consider issue 2.

In matrix notation, nonseparability in a two-dimensional policy space has a deceptively simple definition: $a_{i12} \neq 0$. This means that nonseparability has nothing to do with salience. Salience is determined by the *diagonal* terms in $\mathbf{A}_i$; *nonseparabilities* are captured in the off-diagonal terms of $\mathbf{A}_i$. Larger values of the a_{ij} $(i \neq j)$ represent larger complementarities, or interactions, in the member's evaluation of issues.

To see this, let's expand the formula for $\text{WED}(\mathbf{z}, \mathbf{x}_i)$:

when we have to consider two policies together. For example, a positive interaction is an *increase* in WED, implying a *reduction* in utility. Thus, we will adopt the convention (already outlined intuitively in the preceding chapter) that $a_{12} < 0$ is a positive complementarity, and $a_{12} > 0$ is a negative complementarity, where "positive" and "negative" refer to the effect on the voter's utility.

Recall that, intuitively, nonseparability requires that the voter consider *all* issue positions before choosing *any*. We can summarize the specific impact of the interaction term in three statements about the voter's conditional preferences:

1. If $a_{12} < 0$ (*negative* interaction, implying *positive complementarity*) and the position on issue j is fixed for some reason at $\tilde{x}_j > x_{ij}$, then the conditional ideal on issue k will be *larger* than x_{ik}. If the position on issue j is fixed at $\tilde{x}_j < x_{ij}$, then the conditional ideal on issue k will be *smaller* than x_{ik}.
2. If $a_{12} > 0$ (*positive* interaction, implying *negative complementarity*) and the position on issue j is fixed at $\tilde{x}_j > x_{ij}$, then the conditional ideal on issue k will be *smaller* than x_{ik}. If the position on issue j is fixed at $\tilde{x}_j < x_{ij}$, then the conditional ideal on issue k will be *larger* than x_{ik}.
3. No matter what the form of the interaction, if $\tilde{x}_j = x_{ij}$ (position on issue j is *fixed at the voter's ideal point*), the interaction term is zero. Nonseparability has no impact on choices on issue k: The conditional ideal given z_j is still x_{ik}.

We can derive the conditional preference of the voter (dropping the i subscript for now) for Project 2 by fixing Project 1 at an arbitrary budget we will call $\tilde{x}_1$. (Note: The derivation of the ideal for Project 1, given a fixed budget for Project 2, is analogous and will not be presented.) The voter's *conditional* preference for Project 2, or x_2^C can be written in terms of the WED between $\mathbf{x}$ and the policy vector $[\tilde{x}_1\ x_2^C]^T$. The value of x_2^C that minimizes WED given $\tilde{x}_1$ is unknown, but we can find it in two steps. First, write out the expression for WED:

$$\mathrm{WED}([\tilde{x}_1\ x_2^C]^T, \mathbf{x})$$
$$= \sqrt{a_{11}(\tilde{x}_1 - x_1)^2 + 2a_{12}(\tilde{x}_1 - x_1)(x_2^C - x_2) + a_{22}(x_2^C - x_2)^2} \qquad (4.28)$$

The second step is to find the value of Project 2 budget that conditionally minimizes WED (i.e., maximizes utility). The condition, of course, is that $\tilde{x}_1$ is the (fixed) budget for Project 1. Enelow and Hinich (1984b) performed these two steps and solved for x_2^*, the value of Project 2 that conditionally maximizes the voter's utility:

$$x_2^* | \tilde{x}_1 = x_2 - \left(\frac{a_{12}}{a_{22}}\right)(\tilde{x}_1 - x_1) \tag{4.29}$$

As (4.29) proves, $x_2^* \neq x_2$ whenever preferences are nonseparable and $\tilde{x}_1 \neq x_1$.

There is nothing perverse about this preference rule. In fact, as Enelow and Hinich demonstrate, the preferences in (4.29) are symmetric and single-peaked around the conditional ideal point on issue 2, x_2^*. But $\mathbf{x}$ is still the *package* of proposals the voter most prefers, and it differs from x_2^* on Project 2.

For the rest of this chapter and most of the rest of this book, we will assume separability of preferences. Still, it is crucial to remember that nonseparability is important in modeling real political choice. We will take up the idea of linkages across issues again in Chapters 8 and 9.

Stability of majority rule: The generalized median voter theorem

In Chapter 3, we outlined the problems of defining a Condorcet winner in more than one dimension. We argued that the general condition for stability is the existence of a median in all directions. As the reader will recall, an alternative is a median in all directions if every line drawn through it divides the ideal points of all members so that at least half are on either side of the line. It is important to note that points on the line are counted as belonging to each of the two groups, or half spaces.

This definition is designed for a two-dimensional policy space and is suitable for graphical examples on a printed page, which is after all a two-dimensional surface. How are we to generalize to situations where there are more than two dimensions? Many people find it hard to visualize three dimensions, and we certainly can't represent three-space very well on a flat printed page. Worse, higher dimensions, with 8, 136, or an arbitrary number of policies, defy graphical interpretations. We certainly cannot use "lines" to split groups of ideal points. Mathematicians use a mental construct called a "separating hyperplane" to understand such a splitting of spaces. We will denote the separating hyperplane $\mathcal{H}$. In general, to split a space $\mathcal{H}$ must have *one fewer dimensions* than the space itself.

As we have already seen, in the one-dimensional example $\mathcal{H}$ was of dimension zero (a point). In two-dimensional spaces, $\mathcal{H}$ must be of dimension one (a line). In more complex spaces, with arbitrary dimension n, $\mathcal{H}$ must always be of dimension $n - 1$. For our purposes, $\mathcal{H}$ must have the same property in a space of any dimension: No matter how $\mathcal{H}$ is tilted or rotated, as long as it passes through the median position, $\mathcal{H}$ must divide all ideal points so that (to use the Schwartz, 1986, definition) no majority of ideal lies *strictly* on either side.

Obviously, this condition becomes harder to satisfy in higher dimensions. But the general principle (McKelvey, 1976a, 1976b) is precisely the same in every case: There must be at least $N/2$ ideal points on each side of the separating hyperplane (including in each case points on $\mathcal{H}$ itself).

This principle implies a generalized statement of the median voter theorem, simultaneously accounting for any number of dimensions from one to infinity. For evolving interpretations of this result, see Davis, DeGroot, and Hinich (1972), Kramer (1973), McKelvey (1976a, 1976b, 1986), Schofield (1978a), Slutsky (1979), Cohen and Matthews (1980), and Enelow and Hinich (1983a, 1984b).

Generalized median voter theorem (GMVT)

Assumptions.

(a) Let there be N voters choosing one set of policies in a space of n policy dimensions.
(b) Let preferences be separable and symmetric around each ideal point $\mathbf{x}_i$, for $i = 1, \ldots, N$.
(c) Define a "separating hyperplane" $\mathcal{H}$ of dimension $n - 1$ as a point (when $n = 1$), line (when $n = 2$), plane (when $n = 3$), or hyperplane (when $n > 3$) that divides the N ideal points into two groups. Each group is made up of the ideal points on one side of $\mathcal{H}$, including those ideal points that lie on the surface of $\mathcal{H}$. Then the number of members in each group is N_1 and N_2, respectively.

Theorem. *An alternative* $\mathbf{y}$ *is a median position for the society if and only if, for every* $\mathcal{H}$ *containing* $\mathbf{y}$, $N_1 \geq N/2$, *and* $N_2 \geq N/2$. *The number of* *votes for* **y** *is greater than or equal to the number of votes for any other alternative* **z**.

Interestingly, we have already seen the "proof" of the GMVT.[3] It is the same as that offered in Chapter 2 for the simple MVT in one dimension. The reason is that although the MVT is a special case of the more general theorem, the principle is the same.

Conclusion

This short chapter has been a more formal presentation of the results summarized in Chapter 3. It may be useful, after mastering the mathematical principles discussed here, to reread Chapter 3 as a way of reinforcing the intuition behind the formal discussion.

There have been four main themes in our presentation of the multidimensional model in Chapters 3 and 4:

- A definition of "preference," using distance in space as a metric.
- A discussion of the importance of different priorities, or "salience," of issues in determining preference over two or more issues.
- A discussion of the importance of separability, or the effect of the expected level of one policy on the preferred level of another policy.
- A consideration of the conditions that cause instability of majority rule decision processes when two or more issues are being decided at the same time.

In the next chapter, we consider the instability, and even indeterminacy, of collective choice processes more generally.

EXERCISES

4.1 Suppose $\mathbf{A} = \ell \times \mathbf{I}$, where $0 < \ell < 1$. What is the specific relationship between the SED and the WED measures of distance?

4.2 Suppose committee member i has an ideal point ($\mathbf{x}_i = [10\ 12]^T$) on two budget items and that $\mathbf{A}_i = \begin{bmatrix} 2 & -3 \\ -3 & 1 \end{bmatrix}$. Imagine that for some reason the budget for Project 1 is set at \$20 and that all members take this as given. Is member i's conditional ideal on Project 2, given $\tilde{x}_1 = 20$, larger or smaller than \$12? What is the exact numeric value of $x_{i2}^* \mid \tilde{x}_1$?

4.3 Use the Pythagorean theorem to prove that the SED between two points **z** and **y** in three-dimensional space (dimensions 1, 2, and 3) is:

$$\mathrm{SED}(\mathbf{z}, \mathbf{y}) = \sqrt{(z_1 - y_1)^2 + (z_2 - y_2)^2 + (z_3 - y_3)^2}$$