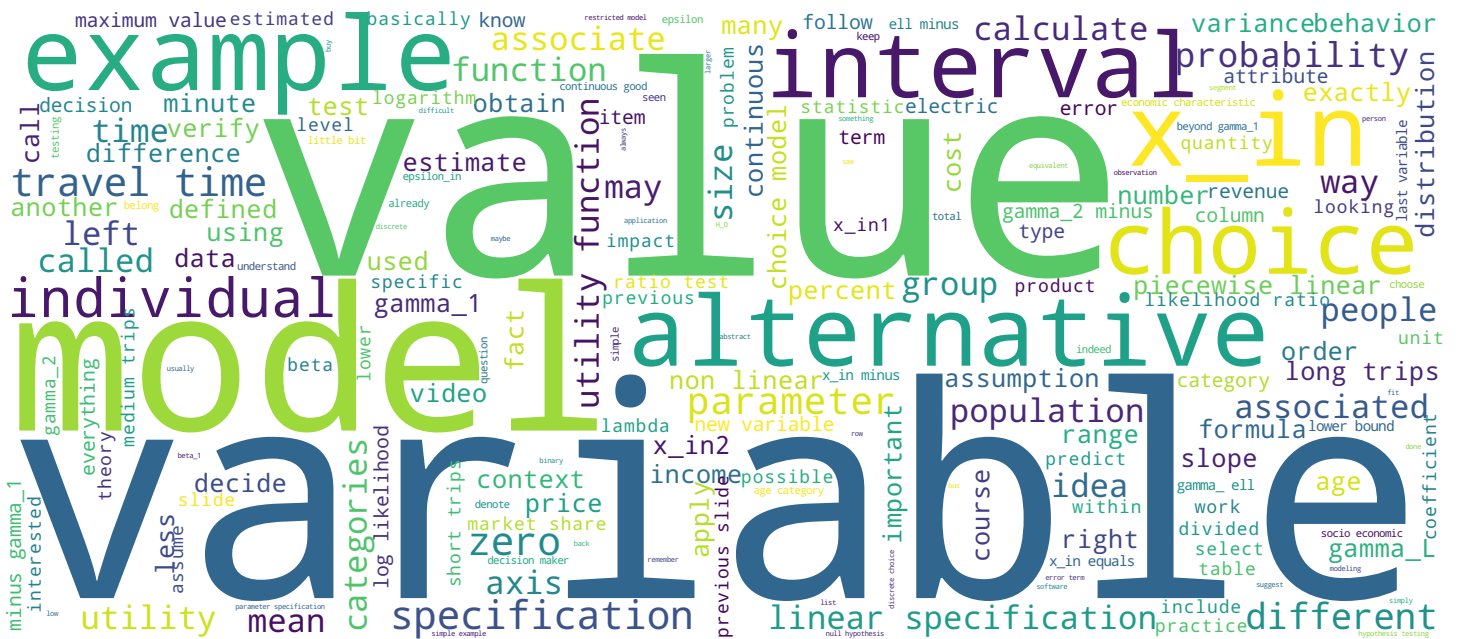


Specification of the deterministic part

Michel Bierlaire

Introduction to choice models

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Search MOOC



Video



EPFL

Nonlinear transformations of the variables

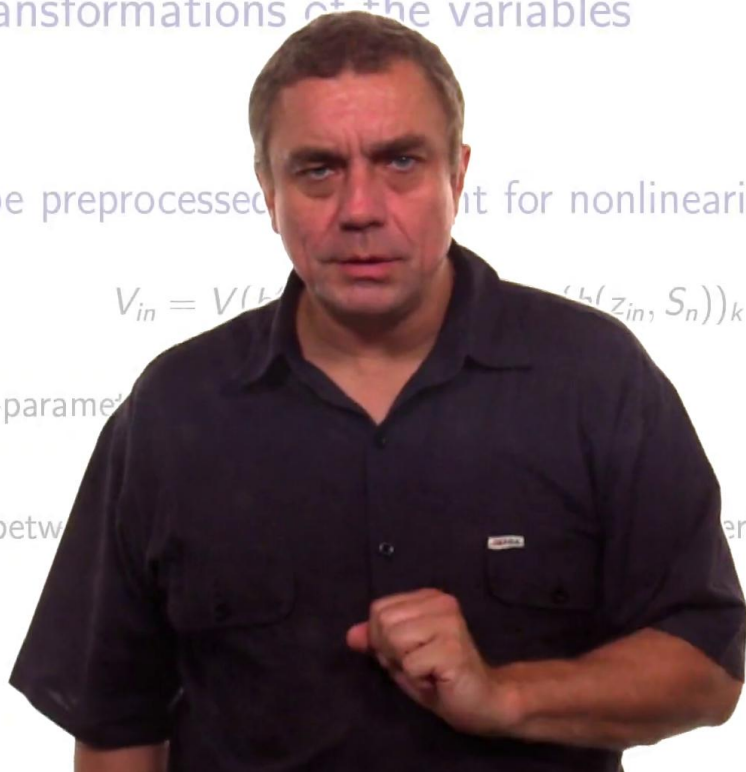
Data can be preprocessed to account for nonlinearities

$$V_{in} = V(h_1(z_{in}, S_n), \dots, h_K(z_{in}, S_n))_k$$

It is linear-in-parameters

Note

Interactions between variables and characteristics are a special case of h



The next specification that we will see is called a piecewise linear specification. Sometimes it's also called a spline. Let's look at what it is.

Notes

Summary



0m 04s

Piecewise linear specification

Again: sensitivity to travel time varies with travel time

- ▶ Log transform is not the only specification
- ▶ Another possibility: split the range of values of the variable
- ▶ Each category is associated with a different coefficient.

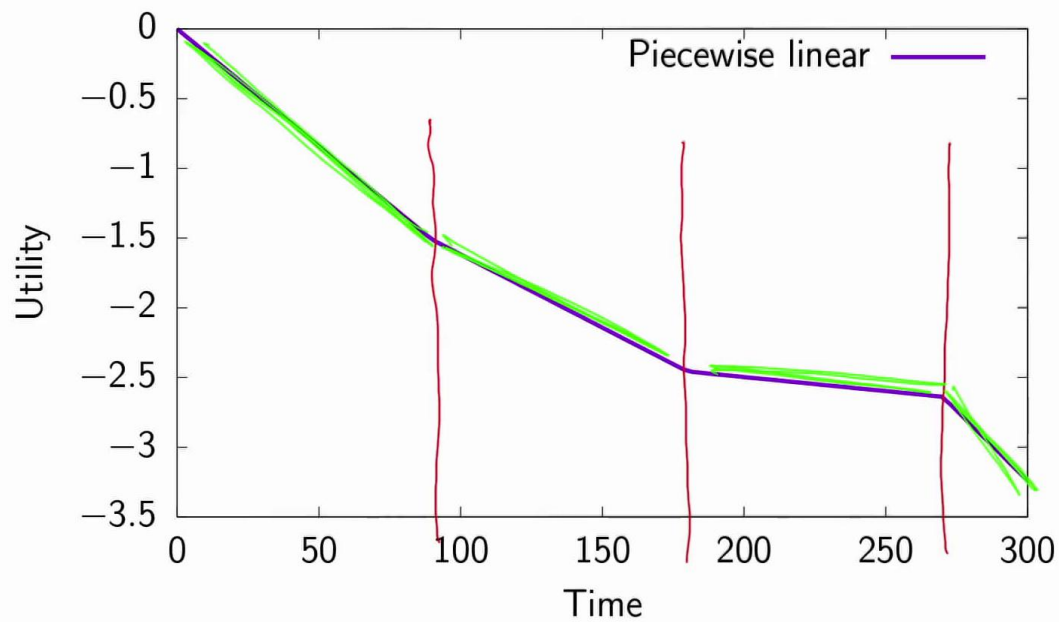
The logarithm is just an example of a possible non linear specification. So if we want to model the fact that the sensitivity to travel time varies with travel time, there are other ways that can be used to do that. Let's see now another example called the piecewise linear specification. The idea is to split the range of values of the variable into categories, and to associate each of these categories to a different coefficient. Let's look at the details.

Notes

Summary



Piecewise linear specification



This is an example of what we want to achieve. Here, the x-axis is travel time. And the y-axis is the utility. And we define categories of travel time by defining breakpoints. Here we have one at 90, another one at 180, and another one at 270. This defines the short trips, medium trips, long trips and very long trips. And as you can see, in this specification, within each category, we have a different slope. We have one slope for short trips, one slope for medium trips, one slope for long trips and one slope for very long trips. This specification is continuous. And it's called piecewise linear because, within each category, the utility function is linear in travel time. But altogether this is a non-linear specification.

Notes

Summary



Piecewise linear specification

Procedure

- ▶ Select breakpoints $\gamma_1 < \gamma_2 < \dots < \gamma_L$
- ▶ Define new variables

Let's now look at the procedure about how to model this mathematically. So we have a variable and we have selected breakpoints that will define the categories. Let's call these break points gamma_1, gamma_2, all the way to gamma_L. In order to write the mathematical specification, we need to define new variables from the existing one.

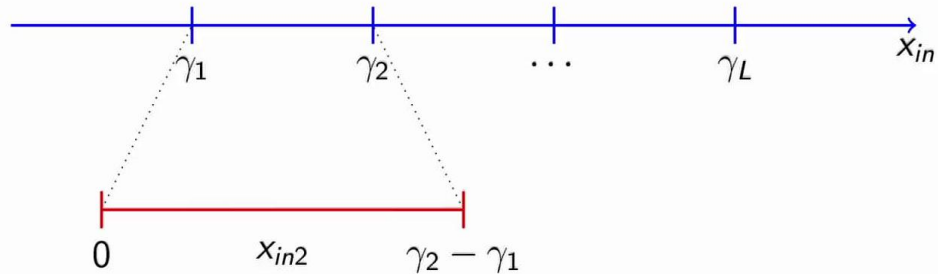
Notes

Summary



2m 03s

Piecewise linear specification



Let's illustrate how this works. So, take the variable x_{in} , and the range of values that this variable can take is represented by this axis here. We have divided this axis into categories using our breakpoints, γ_1 to γ_L , and now we will associate a variable with each of these categories. And this variable will be defined as follows. In this example, we are looking at the second interval, the interval between γ_1 and γ_2 . It is associated with a variable called x_{in2} . The value of this variable will range from zero to γ_2 minus γ_1 . So γ_2 minus γ_1 is the size of the interval. When x_{in} equals γ_1 , the value of x_{in2} is equal to zero. When x_{in} is equal to γ_2 , the value of this variable is equal to γ_2 minus γ_1 . And, in between, the value of x_{in2} varies proportionally to the value of x_{in} . That's the idea. And then we define the same for each of these intervals. So let's now look how it works mathematically, what are the formula for the specification.

Notes

Summary



2m 29s

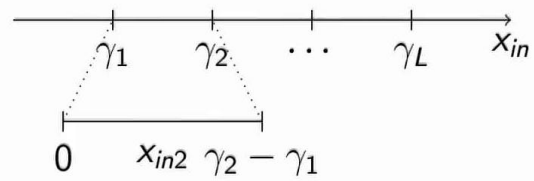
Piecewise linear specification

Formulation

$$x_{in1} = \begin{cases} x_{in} & \text{if } x_{in} < \gamma_1 \\ \gamma_1 & \text{otherwise} \end{cases}$$

$$x_{in\ell} = \begin{cases} 0 & \text{if } x_{in} < \gamma_{\ell-1} \\ x_{in} - \gamma_{\ell-1} & \text{if } \gamma_{\ell-1} \leq x_{in} < \gamma_{\ell} \\ \gamma_{\ell} - \gamma_{\ell-1} & \text{otherwise} \end{cases}$$

$$x_{in(L+1)} = \begin{cases} 0 & \text{if } x_{in} < \gamma_L \\ x_{in} - \gamma_L & \text{otherwise} \end{cases}$$



So the formula is as follows. The first variable x_{in1} is associated with all values lower than γ_1 . And in this case x_{in1} is equal to the value of x_{in} if this value is less than γ_1 , and if it is larger, we assign the value γ_1 to this first variable. Then, within each closed interval, the variable $x_{in\ell}$ is defined as we discussed in the previous slide. So if the value of x_{in} is at the left of the interval, so lower than the lower bound of the interval, the associated variable is equal to zero. If the variable is greater than the upper bound of the interval, the value of the variable is equal to the size of the interval. And, in between, the value of the variable is x_{in} minus $\gamma_{(\ell-1)}$. And you can verify that, when the value of x_{in} is equal to $\gamma_{(\ell-1)}$, so the lower bound of the interval, we'll have zero here. And also when x_{in} is equal to γ_L , this value will be equal to the size of the interval. So this guarantees the continuity of the variable. So the value of this last variable is simply defined as zero if the variable x_{in} is less than γ_L . And x_{in} minus γ_L otherwise. This specification is actually difficult to code in a software, in Biogeme for example.

Notes

Summary



3m 47s

Piecewise linear specification

Equivalent formulations

$$x_{in1} = \begin{cases} x_{in} & \text{if } x_{in} < \gamma_1 \\ \gamma_1 & \text{otherwise} \end{cases} \quad \min(x_{in}, \gamma_1)$$

$$x_{in\ell} = \begin{cases} 0 & \text{if } x_{in} < \gamma_{\ell-1} \\ x_{in} - \gamma_{\ell-1} & \text{if } \gamma_{\ell-1} \leq x_{in} < \gamma_{\ell} \\ \gamma_{\ell} - \gamma_{\ell-1} & \text{otherwise} \end{cases} \quad \max(0, \min(x_{in} - \gamma_{\ell-1}, \gamma_{\ell} - \gamma_{\ell-1}))$$

$$x_{in(L+1)} = \begin{cases} 0 & \text{if } x_{in} < \gamma_L \\ x_{in} - \gamma_L & \text{otherwise} \end{cases} \quad \max(0, x_{in} - \gamma_L)$$

So what we do is that we can use an equivalent formulation which is presented here on the right of the slide. So I let you verify that indeed these two specifications are equivalent.

Notes

Summary



5m 19s

Piecewise linear specification

Equivalent formulations

$$x_{in1} = \begin{cases} x_{in} & \text{if } x_{in} < \gamma_1 \\ \gamma_1 & \text{otherwise} \end{cases} \quad \min(x_{in}, \gamma_1)$$

$$x_{in\ell} = \begin{cases} 0 & \text{if } x_{in} < \gamma_{\ell-1} \\ x_{in} - \gamma_{\ell-1} & \text{if } \gamma_{\ell-1} \leq x_{in} < \gamma_{\ell} \\ \gamma_{\ell} - \gamma_{\ell-1} & \text{otherwise} \end{cases} \quad \min(x_{in} - \gamma_{\ell-1}, \gamma_{\ell} - \gamma_{\ell-1})$$

$$x_{in(L+1)} = \begin{cases} 0 & \text{if } x_{in} < \gamma_L \\ x_{in} - \gamma_L & \text{otherwise} \end{cases}$$

OK but this may look a little bit too abstract. So I suggest that we look now at a concrete example to see how these variables actually coded with specific values.

Notes

Summary



5m 32s

Piecewise linear specification

Examples

$\gamma_1 = 90, \gamma_2 = 180, \gamma_3 = 270.$

$\gamma_1 = 1, \gamma_2 = 5, \gamma_3 = 10.$

x_{in}	50	100	200	300
x_{in1}	50	90	90	90
x_{in2}	0	10	90	90
x_{in3}	0	0	20	90
x_{in4}	0	0	0	30

x_{in}	0.5	4	8	12
x_{in1}	0.5	1	1	1
x_{in2}	0	3	4	4
x_{in3}	0	0	3	5
x_{in4}	0	0	0	2

Utility function

$$V_{in} = \sum_{\ell=1}^L \beta_{\ell} x_{in\ell}$$

On this slide, we have two examples of this piecewise linear specification. On the left, we have the example with the travel time that we illustrated in the previous slides. We take an example where the value of x_{in} , the travel time, can take four values: 50, 100, 200, and 300. If x_{in} is equal to 50, because the value 50 is less than the value of γ_1 , we associate 50 to the first variable and set everything else to 0. In the second column, where x_{in} equals to 100, because it is beyond γ_1 , we set the value of x_{in1} to 90, which is the maximum value that it can take, and the 10 units that are left, we associate them with the second variable x_{in2} . And everything else is equal to zero. When x_{in} is equal to 200, it's beyond γ_1 . So we set the first value here to 90, which is the maximum value that it can take. So what is left is 110, which is beyond the size of the second interval, which is only 90 minutes. So we associate these 90 minutes to the value of x_{in2} . Now we have already set 180 minutes to these variables. So what is left is 20, and 20 can fit in the third interval. So we set x_{in3} to 20. And what is left is zero. And the last example is when we have a value which is beyond γ_3 .

Notes

Summary



5m 43s

Piecewise linear specification

Examples

$\gamma_1 = 90, \gamma_2 = 180, \gamma_3 = 270.$

$\gamma_1 = 1, \gamma_2 = 5, \gamma_3 = 10.$

x_{in}	50	100	200	300
x_{in1}	50	90	90	90
x_{in2}	0	10	90	90
x_{in3}	0	0	20	90
x_{in4}	0	0	0	30

x_{in}	0.5	4	8	12
x_{in1}	0.5	1	1	1
x_{in2}	0	3	4	4
x_{in3}	0	0	3	5
x_{in4}	0	0	0	2

Utility function

$$V_{in} = \sum_{\ell=1}^L \beta_{\ell} x_{in\ell}$$

In this case, we set each of the previous variables to the maximum value that they can take. So in this case all intervals are equals to 90. And then once we have done it, what is left is 30. 300 minus 270. And we associate this to the last variable. So as you can see the sum of all these variables is equal to the value of x_{in} . The example on the right is exactly the same. The difference is that the size of the intervals are different. But the idea is exactly the same. So I let you go through this example on your own to verify that the sum of the $x_{in\ell}$ is equal to the original value of x_{in} . So once you have defined these new variables $x_{in\ell}$ these variables can be included directly in the utility function in a linear-in-parameter specification. And we obtain a piecewise linear specification.

Notes

Summary



7m 14s