### **Estudio R4**

### Explicación del Código

Este código implementa un algoritmo de selección de características utilizando un enfoque de adición hacia atrás (backward addition) que maximiza el error de predicción. Este método es una variante del enfoque de selección hacia adelante del primer código que analizamos y del enfoque de eliminación hacia adelante del tercer código. A continuación, se detalla el funcionamiento del código paso a paso:

### Normalización de los Datos

Al igual que en los códigos anteriores, se centran las variables **X** e **Y** restando sus medias respectivas. Esto es crucial para evitar sesgos en el modelo debido a la escala de las características.

#### Selección de Características

El algoritmo inicialmente considera todas las características disponibles. En cada iteración, intenta añadir una característica y evalúa el modelo con las características actuales. La característica cuya adición resulta en el máximo aumento del error de predicción es considerada la más importante y se añade al conjunto de características.

### **Modelo Polinomial**

Se utiliza un modelo polinomial de orden variable para ajustar los datos y realizar predicciones. Este modelo es similar al utilizado en los códigos anteriores, pero aquí se aplica a un conjunto de características que se va incrementando en cada iteración.

### Cálculo del Error

El error de predicción se calcula utilizando la media de las diferencias absolutas entre las predicciones del modelo y los valores reales de **Y**, elevadas a la potencia de **alpha**. Cuando alpha se establece en 2, lo que corresponde al error cuadrático medio y si es 1 es el error absoluto medio.

# **Comparación con los Códigos Anteriores**

Enfoque de Selección: Mientras que el primer código añadía características de forma iterativa (selección hacia adelante), el segundo código las eliminaba (selección hacia atrás). El tercer código implementaba un enfoque de eliminación hacia adelante, pero seleccionaba la característica cuya eliminación maximizaba el error de predicción. Este cuarto código también implementa un enfoque de adición, pero con una diferencia clave: en lugar de seleccionar la característica que minimiza el error de predicción cuando se añade,

selecciona la característica cuya adición maximiza el error de predicción. Finalmente, se invierte la secuencia de características seleccionadas, por lo que la última característica añadida (que maximiza el error de predicción) se considera la más importante.

### \*Análisis visual de los resultados

Ya sea fijando Alpha=1 o Alpha=2 observo que:

La variable **113** es la que sale primera en la mayoría de cálculos de error **113** y **114** salen en ese orden en las primeras posiciones en todas menos cuando el oren del polinomio es 10

En cuanto al análisis de las variables menos relevantes, vemos que las variables **66 60** y **7**, son las que aparecen en las últimas posiciones.

He subrayado de amarillo las más relevantes y de rojo la que menos. Además, he puesto de color verde la letra si se repite la posición en la mayoría de veces y de naranja si pasa en algunas tablas.

### Estudio de las gráficas:

#### Gráfico de error:

Se muestra un gráfico que representa el error en función del número de características seleccionadas. El eje x representa el número de características seleccionadas, mientras que el eje y representa el valor del error. El primer punto en el gráfico representa el error cuando no se selecciona ninguna característica.

#### Cuando Alpha=1

En cuanto al resultado, la primera "bola" para **Orden poli=2** es **0,23** mientras que los siguientes polinomios de mayor orden, son muy cercanos a **0.21 y 0.2**, mientras que el **orden poli =10** es prácticamente **0.23 aproximadamente**, por lo que el valor de error de la primera variable es prácticamente el mismo.

El error se muestra descendente en todos los casos a medida que aumente el número de variables seleccionadas

Si observamos el recorrido de las variables, es más **constante** cuando el polinomio es de **orden 2 3 4 y 5**, minetras que el polinomio de **orden 10** se muestra más **inestables** los demás polinomios son más **inestables** y se aproxima más a **0**.

En cuanto a las curvaturas, vemos que de forma general tienen forma constante de bajar, a excepción de cuando el orden del **polinomio 10**, que se muestra con mayor curvatura.

# Cuando Alpha=2

En cuanto al resultado, la primera "bola" que se observa en la gráfica, es igual para todos porque es cuando no se selecciona ninguna variable, la siguiente para **Orden poli=2** es **0,08** y vemos que los siguientes polinomios de mayor orden, son muy

cercanos , mientras que el **orden poli =10** es prácticamente **0,1**, por lo que hay menor valor de error con la primera característica

El error se muestra descendente en todos los casos a medida que aumente el número de variables seleccionadas

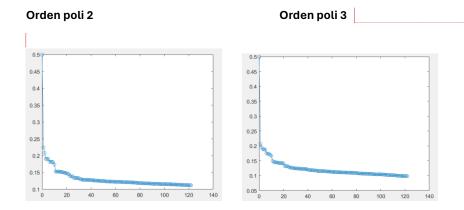
Si observamos el recorrido de las variables, es más **constante** cuando el polinomio es de **orden 2 3 4 y 5**, minetras que el polinomio de **orden 10** se muestra más **inestables** los demás polinomios son más **inestables** y se aproxima más a **0**.

#### \*Análisis Final

La secuencia de errores y su análisis derivativo pueden proporcionar información adicional sobre la estabilidad del proceso de selección de características y la calidad de las características seleccionadas. En resumen, este cuarto código proporciona otra estrategia para la selección de características, que puede ser útil en situaciones donde se desea mantener las características que más contribuyen a la precisión del modelo. La elección entre un enfoque de selección hacia adelante, hacia atrás, de eliminación hacia adelante o de adición hacia atrás, son distintas visiones de realizar estos rankings que puede depender de cuál sea el objetivo del proyecto.

### **GRÁFICAS**

### ALPHA1

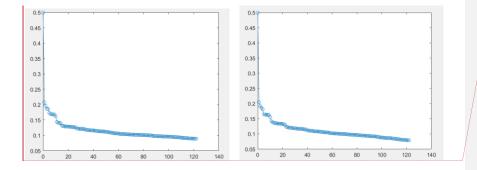


Orden poli 4 Orden poli 5

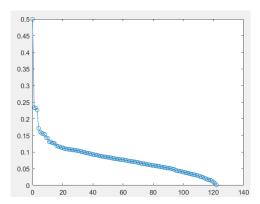
Comentado [CP1]: orden poli 2

Comentado [CP2]: orden poli 4

Comentado [CP3]: orden poli 5

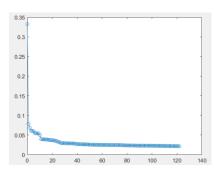


# Orden poli 10

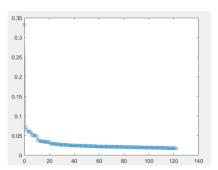


### **ALPHA2**

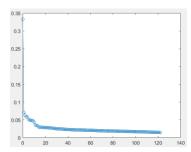
# Orden poli 2



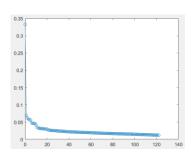
# Orden poli 3



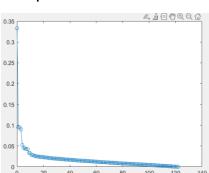
# Orden poli 4



# Orden poli 5

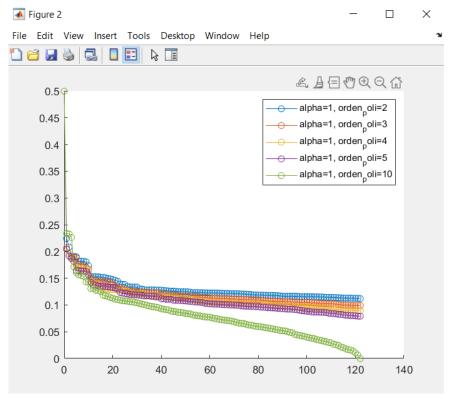


# Orden poli 10

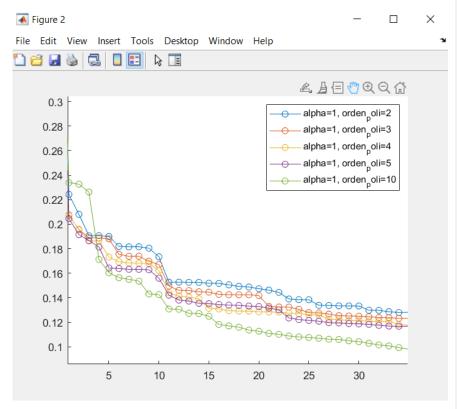


# Gráficas comparativas

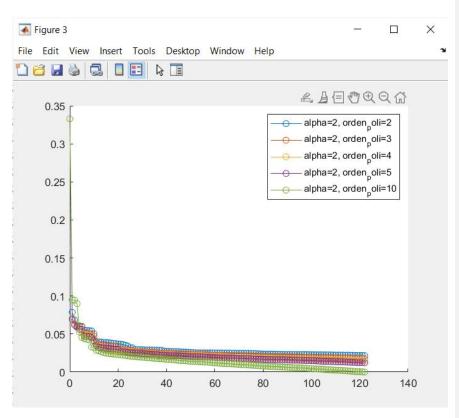
# **CUANDO ALPHA 1**



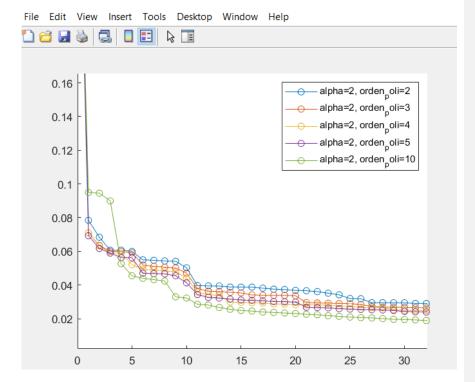
Zoom:



**CUANDO ALPHA 2:** 



Zoom:



# **TABLA DE DATOS:**

```
{'Alpha'} {'Orden Poli'} {'Error'
                                        } {'Solución'
 {[ 1]} {[
                 2]} {[0.2243 0.2080 ... ]} {[113 114 115 1 ... ]}
                 3]} {[0.2073 0.1956 ... ]} {[113 114 1 115 ... ]}
  {[ 1]} {[
 {[ 1]} {[
                 4]} {[0.2076 0.1954 ... ]} {[113 114 1 115 ... ]}
 {[ 1]} {[
                 5]} \{[0.2046\ 0.1915\ ...\ ]\}\ \{[\ 113\ 114\ 50\ 1\ ...\ ]\}
  {[ 1]} {[
                10]\} \ \ \{[0.2338\ 0.2326\ ...\ ]\} \ \ \{[\ 115\ 1\ 51\ 113\ ...\ ]\}
{'Alpha'} {'Orden Poli'} {'Error'
                                        } {'Solución' }
 {[ 2]} {[
                 2]} {[0.0784 0.0684 ... ]} {[113 114 1 115 ... ]}
  {[ 2]} {[
                 3]} \{[0.0710\ 0.0635\ ...\ ]\}\ \{[113\ 114\ 1\ 115\ ...\ ]\}
  {[ 2]} {[
                 4]} {[0.0709 0.0632 ... ]} {[113 114 1 115 ... ]}
  {[ 2]} {[
                 5]} {[0.0693 0.0617 ... ]} {[ 113 114 50 1 ... ]}
 {[ 2]} {[
                10]} {[0.0950 0.0945 ... ]} {[ 1 115 51 113 ... ]}
```

# Tablas de los resultados

# Cuando Alpha =1

# Orden del Polinomio 2

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	113	<mark>114</mark>	115	1	50	116	2	51	14	20	119	12	18	8	13
%															
2	9	21	117	15	11	88	23	87	122	6	4	3	19	16	27
%															
3	56	7	85	29	86	25	38	118	53	31	33	121	47	10	42
%															
4	37	40	39	17	26	41	44	43	45	46	65	64	48	49	69
%															
5	89	112	5	98	108	95	91	90	94	93	92	59	22	72	96
%															
6	97	57	58	67	28	75	102	70	35	62	36	61	83	100	80
%															
7	99	81	109	24	32	63	52	30	54	105	82	77	101	110	74
%															
8	79	78	34	107	76	<mark>60</mark>	106	66	68	103	104	71	111	73	84
%															
9	55	120													

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	<mark>113</mark>	<mark>114</mark>	1	115	50	116	2	51	20	14	12	119	8	13	21
%															
2	22	18	117	9	15	4	6	122	88	87	23	11	19	118	3
%															
3	25	16	85	29	38	44	37	58	39	56	27	42	41	40	17
%															

4	31	47	43	46	45	53	111	99	68	48	97	49	69	86	33
%															
5	90	98	64	95	61	59	62	92	108	91	89	120	77	5	75
%															
6	74	109	76	26	35	72	36	57	93	65	82	83	70	102	52
%															
7	63	10	121	80	112	28	81	54	105	96	55	110	30	32	79
%															
8	78	101	107	34	106	73	103	104	100	67	84	71	24	94	<mark>66</mark>
%															
9	60	7													

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	<mark>113</mark>	<mark>114</mark>	1	115	116	50	2	51	14	20	119	12	16	18	4
%															
2	21	13	6	9	117	15	8	19	23	88	122	87	37	3	118
%															
3	17	38	11	56	52	63	22	85	31	29	65	25	27	69	101
%															
4	103	47	45	5	41	28	86	53	39	43	42	40	44	120	109
%															
5	33	35	48	46	49	26	95	71	108	76	112	10	64	55	60
%															
6	98	81	106	62	61	90	67	77	75	36	92	58	96	59	89
%															
7	57	74	80	68	72	93	91	82	121	111	83	105	99	102	110
%															
8	100	107	73	104	32	30	54	70	34	97	79	84	24	78	94
%															
9	7	<mark>66</mark>													

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	113	<mark>114</mark>	50	1	116	115	2	51	14	20	12	119	8	22	21
%															
2	13	9	18	15	117	19	6	4	16	3	17	52	63	65	38
%															
3	29	85	87	37	118	31	122	88	23	56	101	69	11	25	47
%															
4	27	81	109	26	39	40	41	42	120	5	121	53	43	86	33
%															
5	28	44	95	45	96	46	59	71	108	76	64	48	49	10	55
%															
6	92	60	98	62	97	58	90	89	57	83	111	67	35	68	75
%															
7	82	61	91	36	105	93	99	30	32	102	70	34	80	78	112
%															
8	106	77	72	24	54	74	110	107	84	94	79	100	103	104	73
%															
9	66	7													

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	115	1	51	<mark>113</mark>	50	<mark>114</mark>	2	116	24	14	119	20	12	8	15
%															
2	4	16	3	18	21	19	117	9	22	13	6	64	65	59	17
%															
3	97	38	41	52	118	11	85	32	66	47	109	62	39	43	112
%															
4	91	56	95	67	37	61	69	31	5	92	29	45	27	49	48
%															
5	40	101	42	28	44	46	33	82	26	35	36	55	58	107	89
%															

6	90	57	71	93	73	120	122	87	53	88	25	23	98	30	75
%															
7	83	96	86	76	108	121	102	60	63	54	81	10	80	34	70
%															
8	99	78	68	79	105	77	110	94	111	74	104	103	106	84	72
%															
9	100	7													

# Cuando Alpha =2

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	113	<mark>114</mark>	1	115	50	116	2	51	14	20	119	12	8	13	9
%															
2	19	18	11	21	117	15	88	23	122	87	6	4	3	38	16
%															
3	85	22	7	25	29	118	86	27	56	53	31	33	26	47	39
%															
4	37	17	42	40	44	41	43	45	46	65	64	112	48	49	28
%															
5	70	89	5	95	74	61	99	62	98	90	92	91	77	75	57
%															
6	59	76	58	68	72	10	80	79	35	78	73	109	83	101	81
%															
7	63	36	93	96	32	82	110	107	69	102	30	97	105	34	108
%															
8	94	54	121	24	100	67	55	71	52	84	<mark>60</mark>	<mark>66</mark>	106	103	111
%															
9	104	120													

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	<mark>113</mark>	<mark>114</mark>	1	115	50	116	2	51	14	20	119	12	8	13	21
%															
2	22	18	117	9	15	4	19	6	11	122	88	87	23	3	118
%															
3	85	16	25	29	27	56	33	31	86	7	53	38	44	5	47
%															
4	102	90	58	77	26	39	28	40	97	37	17	42	43	41	45
%															
5	46	48	49	112	110	74	62	64	35	76	75	95	61	89	99
%															
6	91	92	59	65	57	70	72	93	36	108	80	68	82	98	105
%															
7	78	79	73	63	83	81	10	32	94	96	69	111	107	34	55
%															
8	84	30	71	100	54	24	121	106	104	103	52	120	67	101	109
%															
9	66	60													

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	<mark>113</mark>	<mark>114</mark>	1	115	116	50	2	51	14	20	119	12	16	4	22
%															
2	21	8	13	9	19	6	117	15	3	18	17	52	122	87	88
%															
3	85	38	11	7	23	56	29	25	27	86	118	37	39	44	63
%															
4	47	98	61	111	99	68	96	62	97	90	91	57	92	95	26
%															
5	53	31	41	45	43	33	40	42	48	49	46	28	24	30	5
%															

6	108	89	77	32	75	72	76	55	74	73	35	70	94	82	58
%															
7	93	80	78	36	59	65	64	112	110	105	54	102	83	81	121
%															
8	10	79	69	84	34	107	100	71	67	106	104	103	101	120	60
%															
9	66	109													

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	113	<mark>114</mark>	50	1	115	116	2	51	14	20	12	119	18	22	19
%															
2	21	8	13	9	117	4	6	17	52	3	38	37	118	88	87
%															
3	122	85	15	16	11	23	56	29	25	27	86	63	94	62	112
%															
4	59	91	65	64	69	99	93	92	109	90	98	58	28	47	41
%															
5	39	53	5	33	31	40	42	48	43	44	45	46	49	95	108
%															
6	76	26	10	121	111	55	35	120	54	57	68	72	30	75	32
%															
7	89	77	36	74	61	96	67	80	70	102	83	81	82	34	24
%															
8	105	97	73	78	110	100	101	7	79	107	106	104	103	84	71
%															
9	60	<mark>66</mark>													

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	115	1	51	<mark>113</mark>	50	114	116	2	4	14	16	20	12	119	18
%															
2	10	15	117	19	22	21	8	9	13	11	86	65	121	38	104
%															
3	122	85	118	112	31	94	62	3	27	59	91	26	37	63	69
%															
4	6	52	23	57	66	80	29	25	88	56	5	87	28	110	108
%															
5	53	71	111	96	90	102	41	39	40	42	64	120	76	33	97
%															
6	58	105	55	17	89	35	95	93	24	47	46	92	82	61	36
%															
7	75	45	43	67	44	48	49	60	98	81	84	106	83	54	103
%															
8	32	30	34	100	68	72	107	7	78	77	70	99	79	73	101
%															
9	109	74													