a) **Modelo Lineal**

En este caso, supondremos que se ajustará un modelo lineal suponiendo que el sistema real es lineal con ruido blanco gaussiano aditivo, es decir,

Luego, se propone un modelo lineal para llevar a cabo la predicción a 1 paso, de modo tal que:

Predicción a 1 paso:

Este modelo no considera un valor constante o bias dado el supuesto que el sistema es lineal con ruido blanco aditivo. En caso que se sospechara que existe un bias o tendencia (trend) en el sistema, se puede agregar otro vector de unos a la matriz de regresores (o matriz de información).

Para llevar a cabo la estimación de los parámetros del modelo se utilizó la técnica de mínimos cuadrados, es decir:

En que es el vector de parámetros y es la matriz de regresores con los valores de las n muestras ordenados por filas del conjunto de entrenamiento, es decir,

Los valores que se obtuvieron de los parámetros fueron los siguientes:

En las figuras siguientes (Figura 1, 2 y 3), se muestran las gráficas de la predicción a un paso en cada conjunto de datos, a saber, conjunto de entrenamiento, prueba y validación. Esto es solo a modo de comparación puesto que finalmente es lo relevante de nuestro modelo, es el comportamiento o bondad de ajuste en el conjunto de validación. Evidentemente que los datos de los parámetros usados en las predicciones, para todos los conjuntos, son aquellos obtenidos con el conjunto de entrenamiento, sin embargo, la matriz de represores y salidas cambian dependiendo del conjunto que se esté usando para predicción.



Figura 1: Predicción a 1 paso en conjunto de entrenamiento



Figura 2: Predicción a 1 paso en conjunto de prueba



Figura 3: Predicción a 1 paso en conjunto de validación

A continuación, en la Tabla 1, se presentan las métricas de bondad del ajuste o errores en los diversos conjuntos de datos, a saber, conjunto de datos de entrenamiento, prueba o test y validación.

Tabla 1: Errores o Métricas de bondad de ajuste a 1 paso

|  |  |  |  |
| --- | --- | --- | --- |
|  | Conjunto Entrenamiento | Conjunto de Prueba | Conjunto de Validación |
| RMSE | 0,0869 | 0,0559 | 0,0954 |
| MAPE | 162.2793 | 220,0863 | 238,2150 |
| MAE | 0,3569 | 0,3465 | 0,6020 |

Es interesante notar que los ajustes del modelo, en todos los conjuntos, son bastante comparables entre sí, es decir, en general, los errores son similares en cuanto a RMSE y MAE. Esto también se puede apreciar al observar las figuras 1,2 y 3 respectivamente. Si bien, el conjunto de entrenamiento presenta el menor MAPE, a su vez, el conjunto de entrenamiento, presenta el menor RMSE.

Finalmente, para efectos de apreciar la robustez de este modelo, llevaremos a cabo la predicción a 8 y 16 pasos respectivamente, en el conjunto de validación del modelo lineal, para luego compararlos con los demás modelos, a saber, modelo de Takagi y Sugeno y Modelo Neuronal para el mismo conjunto de validación.



Figura 4: Predicción a 1 paso en el conjunto de validación para el Modelo Lineal.

Figura 5: Predicción a 8 pasos en el conjunto de validación para el Modelo Lineal.



Figura 6: Predicción a 16 pasos en el conjunto de validación para el Modelo Lineal.

**b) Construcción de Intervalos para el modelo lineal**

En este caso, se pueden definir 2 modelos adicionales que den cuenta de un intervalo que contenga un porcentaje de las muestras al interior del mismo. A saber, podríamos determinar un modelo por arriba (y\_u) y uno por abajo (y\_l) del modelo de valor esperado y\_val, con los datos del conjunto de validación, determinado en la sección anterior.

Estos modelos se podrían construir considerando la adición de la varianza de las salidas del conjunto de validación entre el valor real y el estimado, es decir, el intervalo tendría la siguiente expresión para ambos modelos:

En que α es un parámetro de ajuste que tiene relación con el porcentaje de cobertura que tendría el intervalo, Var() es el operador varianza de las salidas reales y estimadas en el conjunto de validación y es el modelo obtenido de valor esperado visto en la sección anterior, es decir,

Así entonces, podemos determinar los modelos por arriba y por abajo del modelo clásico de valor esperado, los cuales nos definen el intervalo.

A continuación, en la Figura 7, se muestran las gráficas de las 3 curvas o modelos para la predicción a un paso con α=0,8.

Figura 7: Curvas para Intervalo del modelo lineal considerando predicción a 1 paso y α=0,8.

En la Figura 8, se muestra un zoom de una porción del gráfico para apreciar mejor el intervalo.

Figura 8: Porción del gráfico de la Figura 7.

Ahora bien, podemos ir ajustando el valor de **α** para aumentar o disminuir el porcentaje de cobertura de las muestras de validación.

En la siguiente tabla, se muestran las métricas PICP y PINAW las cuales miden el porcentaje de cobertura y ancho del intervalo respectivamente, para diferentes valores de α de la predicción a 1 paso.

Tabla 2: Métricas de porcentaje de cobertura y ancho del intervalo para diferentes valores de α.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Parámetro α | 0,5 | 0,8 | 0,9 | 1 | 1,5 | 2 |
| PICP (%) | 32,2034 | 52,5424 | 58,7288 | 65,3390 | 88,5593 | 96,1017 |
| PINAW | 11,687 | 18,6992 | 21,0366 | 23,7339 | 35,0609 | 46,7479 |

A continuación, en el siguiente gráfico, se muestra el caso en que el parámetro α se aumenta a 1,5.



Figura 9: Curvas para Intervalo del modelo lineal considerando predicción a 1 paso y α=1,5.



Figura 10: Porción del gráfico de la Figura 9.

Es evidente, que los datos de salida de la curva promedio (color azul) están casi completamente contenidos en el intervalo (entre curva de color rojo y verde respectivamente), de hecho, está en un 88,56% contenida en el intervalo.

Ahora bien, en el caso de la construcción del intervalo de predicción a 8 y 16 pasos, se procede de manera análoga al caso de 1 paso, con la salvedad que se debe usar el modelo de 8 pasos y 16 pasos respectivamente.

A continuación, en las figuras 11 y 12 se muestran los intervalos para el caso de 8 y 16 pasos considerando un parámetro α=0.8, para efectos de comparación con el gráfico de la Figura 7.

Figura 11: Curvas para Intervalo del modelo lineal considerando predicción a 8 pasos y α=0,8.



Figura 12: Curvas para Intervalo del modelo lineal considerando predicción a 16 pasos y α=0,8.

Tabla 3: Métricas de porcentaje de cobertura y ancho del intervalo para diferentes valores de α=0,8 considerando intervalos de predicción a 1, 8 y 16 pasos respectivamente.

|  |  |  |  |
| --- | --- | --- | --- |
| **Parámetro α** | **0,8** | | |
|  | 1 paso | 8 pasos | 16 pasos |
| PICP (%) | 52,5424 | 52,7966 | 52,7966 |
| PINAW | 18,6992 | 18,4244 | 18,7348 |

Analizando las figuras anteriores y de la Tabla 3, se puede concluir que las métricas de bondad para los intervalos son bastante similares. Por una parte, entendemos que esto se debe a que como el número de muestras del conjunto es muy elevado (1.200 muestras aproximadamente), respecto a las predicciones que se piden (1, 8 y 16 pasos), es más o menos evidente que no exista diferencia en los porcentajes de cobertura. La diferencia podría notarse en caso que los intervalos de predicción fueran a más muestras, por ejemplo, a 200 muestras en adelante. Otro factor que favorece la similitud de las métricas anteriores, es el hecho que la entrada (señal APRBS) es bastante aleatoria y por tanto el ajuste no se ve más favorecido en un conjunto que en otro (conjunto de prueba o validación, por ejemplo). Es decir, los parámetros de ajuste que se obtienen son similares, no importando que conjunto que usemos para entrenamiento. Es evidente que el caso del conjunto de entrenamiento, lo favorece el hecho de que cuenta con más datos, pero dada la aleatoriedad de la señal de entrada, este factor no es tan incidente a la hora de comparar las métricas.

Esta tabla debe ir al final de la parte b)

A continuación, en la Tabla N°xx, se muestran las diversas métricas de bondad de predicción para los diferentes modelos analizados en este informe (Lineal, Takagi y Sugeno y Neuronal) usando los datos de validación.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Modelo Lineal | | | Modelo de Takagi y Sugeno | | | | Modelo Neuronal | | |
|  | 1 paso | 8 pasos | 16 pasos | 1 paso | 8 pasos | 16 pasos | 1 paso | | 8 pasos | 16 pasos |
| RMSE | 0,0954 | 0,0789 | 0,0964 |  |  |  |  | |  |  |
| MAPE | 238,2150 | 226,6030 | 235,3357 |  |  |  |  | |  |  |
| MAE | 0,6020 | 0,5805 | 0,5944 |  |  |  |  | |  |  |

Tabla xx: Errores ó Métricas de bondad de ajuste para predicción resumida en los conjuntos de validación para los tres modelos obtenidos.