Predicción de energía eléctrica generada por un sistema electromecánico de molinos

Carrera de matemática UMSA

Univ. Ticona Coaquira Franklin Josue

Análisis matricial

Mayo 2021

ÍNDICE GENERAL

CAPÍTULO

1

OBJETIVOS DEL PROYECTO

1.1. Objetivo general

Diseñar un predictor de energía eléctrica producida por un sistema de energía eólico.

1.2. Objetivos específicos

- Analizar la importancia de la potencia eléctrica en la industria.
- Desarrollar la teoría de regresión lineal en varias variables.
- Implementar un predictor basado en una regresión lineal en Python.

CAPÍTULO

2

ENERGÍA ELÉCTRICA

2.1. Sistemas de generación de energía eléctrica

El funcionamiento de un sistema de generación de energía electro-mecánico está basado en la ley de Faraday la cual indica lo siguiente:

Ley 2.1 (Ley de Faraday)**.** La f.e.m. inducida es igual a menos la derivada sustancial (total) del flujo con respecto al tiempo.

$$\varepsilon = -\frac{d\Phi}{dt}$$

Gracias a esta ley es que gracias al cambio de un flujo magnético (por lo general se logra a través del movimiento de imanes) es que se puede generar un voltaje (f.e.m) el cual puede ser almacenado para su posterior distribución.

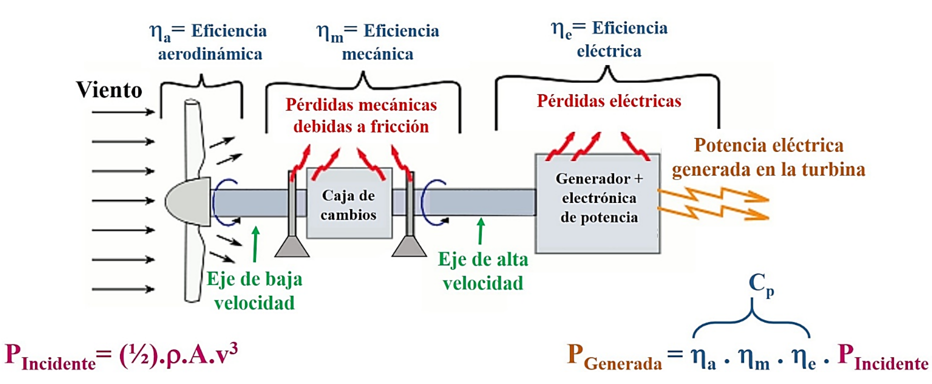


Figura 2.1: Sistema electromecánico de un molino

Recordemos que cuando un voltaje está presente en un circuito eléctrico-electrónico este modifica la dinámica de las corrientes y voltajes de todos los componentes eléctricos presentes, en el caso del molino, el voltaje generado (explicado por la ley de Faraday) modificaría la dinámica de todo circuito conectado a este, por ejemplo en la imagen 2.2, el Generador+electrónica de potencia tendría una dinámica muy compleja de analizar (sistemas de ecuaciones diferenciales no lineales), por esta razón no suele emplearse directamente a la corriente o voltaje como parámetro que cuantifique la energía producida por un generador eléctrico, sino suele emplearse el concepto de potencia:

Definición 2.1. *(Potencia instantánea)* La potencia instantánea $p(t)$ absorbida por un elemento es el producto de la tensión instantánea $v(t)$ en las terminales del elemento y la corriente instantánea $i(t)$ a través de él. Suponiendo la convecnión pasiva de los signos:

$$p(t) = v(t)i(t) \tag{2.1}$$

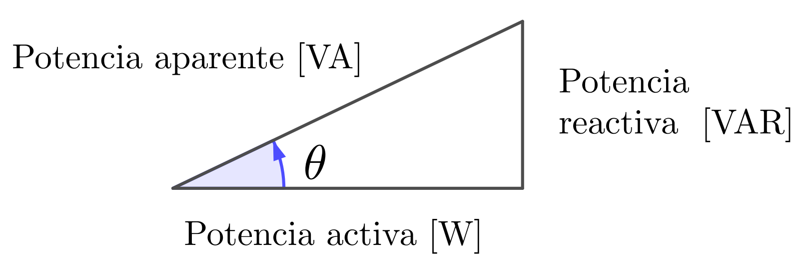
Las ventajas de emplear el concepto de potencia son:

- Si se considera que todos los voltajes de entrada de los circuitos son senoidales (voltajes industriales), entonces la potencia total del circuito puede calcularse como la suma algebraica de las potencias individuales de los componentes.

  $$P_{suministrada} = P_1 + P_2 + \cdots + P_n \tag{2.2}$$

  Donde P puede ser la potencia real, potencia reactiva y potencia compleja del circuito.
- Es de fácil medición.
- Las unidades de costo energético $[KWh]$ son altamente conocidas por profesionales de otras áreas.
- Máquinas eléctricas empleadas en la industria (por ejemplo motores eléctricos) suelen especificar la potencia necesaria para su funcionamiento.

Cuando se trabaja con potencia fasorial (voltaje y corriente senoidal), puede dibujarse el siguiente diagrama de potencias,este diagrama presenta una relación geométrica de los 3 tipos de potencia existentes:



$$f.p. = \text{factor de potencia} = \cos\theta$$

Ahora veamos un ejemplo de aplicación donde se aplica directamente la definición de potencia activa:

2.2. Ejemplo, costo total de energía de una empresa productora de Yeso

Una industria productora de Yeso consume 200[MWh] al mes. Si su demanda máxima es de 1600[KW], calcule la tarifa de electricidad si:

- Cargo de demanda: \$5,00 al mes por KW de demanda facturable.
- Cargo de energía: 8 centavos por KWh para los primeros 50000[KWh], 5 centavos por [KWh] para la energía restante.

Hallemos el cargo de demanda:

$$C_d = \$5{,}00 \times 1600 = \$8000 \tag{2.3}$$

Cargo de energía por los primeros 50000[KWh]:

$$C_p = \$0{,}08 \times 50000 = \$4000 \tag{2.4}$$

Energía restante 200000[KWh] $-$ 50000[KWh] $=$ 150000[kW], cargo de la energía restante:

$$C_r = \$0{,}05 \times 150000 = \$7500 \tag{2.5}$$

De este modo:

$$C_{total-mensual} = C_d + C_p + C_r = \$19500 \tag{2.6}$$

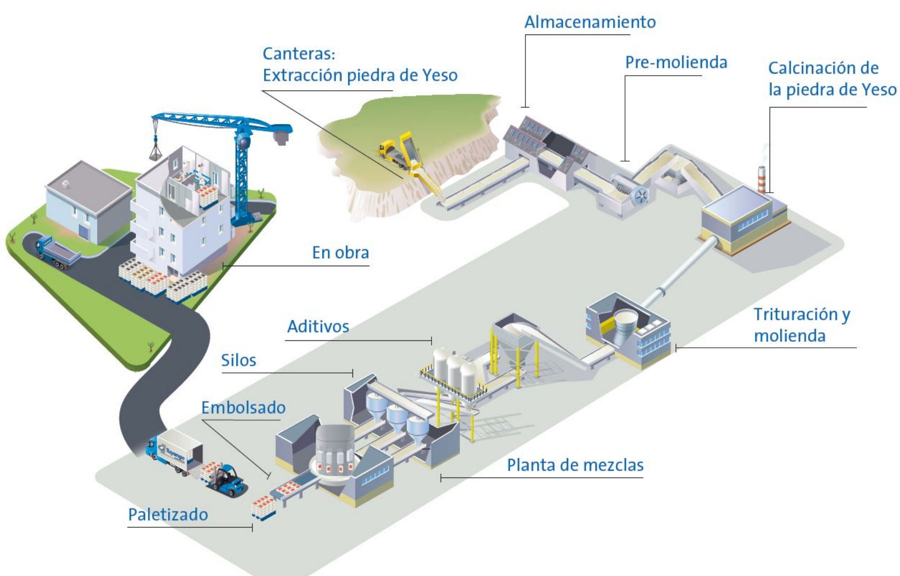


Figura 2.2: Empresa productora de Yeso

Es importante aclarar que existen ciertas máquinas de una empresa (elementos principales de un sistema de producción) que requieren energía en cada momento, caso contrario la empresa podría sufrir pérdidas económicas (por ejemplo, en la empresa de producción de Yeso la máquina principal es el horno rotario que calcina el Yeso molido-húmedo).

CAPÍTULO

3

PREDICCIÓN DE ENERGÍA ELÉCTRICA

Recordando el anterior capítulo, la potencia suministrada a ciertas máquinas debe ser siempre la suficiente (sino pueden producirse pérdidas económicas y por tanto podría caer una demanda en la empresa productora de energía),además, como se vio en la primera parte, es bastante complejo emplear los modelos matemáticos del sistema electromecánico del molino para establecer una relación directa entre la potencia generada y condiciones climáticas, por ejemplo la ecuación que gobierna la dinámica de los fluidos (en este caso aire) viene dada por la ecuación de Navier Stokes:

$$\rho\frac{du_i}{dt} = -\frac{\partial p}{\partial x_i} + \sum_k \frac{\partial}{\partial x_k}\left(\mu\left(\frac{\partial u_i}{\partial x_k} + \frac{\partial u_k}{\partial x_i}\right)\right) + \frac{\partial}{\partial x_i}\left(\lambda \sum_k \frac{\partial u_k}{\partial x_k}\right) + f_i \quad \text{para} \quad i \in \{1,2,3\} \tag{3.1}$$

Dada la dificultad de modelar el molino (sistema electrómecanico que es inducido por un fluído) y con el fin de conocer si es que se cumplirá con la cota mínima de energía a generar, se vio la necesidad de diseñar un predictor de energía electríca basada en una regresión lineal, de esta forma siempre se tendría un control en la potencia generada por la empresa.
Antes de la implementación en Python, desarrollemos la teoría necesaria para entender el funcionamiento del predictor.

3.1. Regresión lineal

Modelo 3.1. Supóngase que desea establecerse una relación matemática entre varias variables observadas $\{x_1, x_2, \ldots, x_k\}$ y una variable de salida $\{y\}$, para tal propósito suponemos que esta relación viene dada por el siguiente modelo lineal:

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k + \epsilon \tag{3.2}$$

Ahora el problema radica en encontrar los β_i que minimicen el error del modelo lineal establecido. Para ello podemos suponer que:

- $E(\epsilon) = 0$
- $Var(\epsilon) = \sigma^2$

Y que las variables aleatorias del vector aleatorio ϵ no están correlacionadas.
Bajo estas suposiciones podemos formular el siguiente problema de optimización:

$$\begin{aligned} \text{minimizar} \quad & S(\beta) = \sum_{i=1}^{n} \epsilon_i^2 = \epsilon^T \epsilon \\ \text{Sujeto a} \quad & \mathbf{y}_{n\times 1} = \mathbf{X}_{(n+1)\times(k+1)} \beta_{k\times 1} + \epsilon_{n\times 1} \end{aligned}$$

Obteniendo que:

Teorema 3.1. La solución óptima β que minimiza el anterior problema de optimización es:

$$\hat{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y} \tag{3.3}$$

Demostración. Haciendo manipulaciones matriciales:

$$S(\beta) = \epsilon^T\epsilon = (\mathbf{y} - \mathbf{X}\beta)^T(\mathbf{y} - \mathbf{X}\beta) \tag{3.4}$$

$$S(\beta) = \mathbf{y}^T\mathbf{y} - 2\beta^T\mathbf{X}^T\mathbf{y} + \beta^T\mathbf{X}^T\mathbf{X}\beta \tag{3.5}$$

Mostremos que $S(\beta)$ es convexa, empleando la condición de 2do orden para S, ya que S es dos veces diferenciable, con un dominio convexo, además:

$$\nabla^2 S(\beta) = 2\mathbf{X}^T\mathbf{X} \succeq 0 \quad \forall \beta \in dom(S) \tag{3.6}$$

Entonces queda probado que S es convexa, ahora hallemos un punto crítico de S:

$$\frac{\partial S(\beta)}{\partial \beta} = -2\mathbf{X}^T\mathbf{y} + 2\mathbf{X}^T\mathbf{X}\beta = \mathbf{0} \tag{3.7}$$

Entonces:

$$\beta = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y} \tag{3.8}$$

Como S es convexa este punto crítico es igual al mínimo global de S, por tanto:

$$\hat{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y} \tag{3.9}$$

$\square$

Además si añadimos los siguientes supuestos:

- $E(\mathbf{y}) = \mathbf{X}\beta$
- $Cov(\mathbf{y}) = \sigma^2 I$

Podemos mostrar los siguientes resultados:

- $\hat{\beta}$ es estimador no sesgado.
- $Cov(\hat{\beta}) = \sigma^2(\mathbf{X}^T\mathbf{X})^{-1}$
- Los estimadores por mínimos cuadrados $\hat{B}_i$, $i = 0, \ldots, k$ tienen varianza mínima entre todos los estimadores lineales no sesgadoz de β.
- El mejor estimador no sesgado de $a'\beta$ es $a\hat{\beta}$, donde $\hat{\beta}$ es el mejor estimador no sesgado de β.
- Si $\mathbf{x} = (1, x_1, \ldots, x_k)^T$ y $z = (1, c_1x_1, \ldots, c_kx_k)^T$, entonces:

$$\hat{\mathbf{y}} = \hat{\beta}^T\mathbf{x} = \hat{\beta}_z^T\mathbf{z} \tag{3.10}$$

  donde $\hat{\beta}_z$ es el estimador de la regresión de y sobre z.

También tenemos el siguiente resultado para estimar σ^2:

Teorema 3.2. *Definiendo s^2 como $s^2 = \frac{1}{n-k-1}(\boldsymbol{y}^T\boldsymbol{y} - \hat{\beta}^T\boldsymbol{X}^T\boldsymbol{y})$, si $E(\boldsymbol{y}) = \boldsymbol{X}\beta$ y $Cov(\boldsymbol{y}) = \sigma^2 I$, entonces:*

$$E(s^2) = \sigma^2 \tag{3.11}$$

Finalmente podemos mostrar los siguientes resultados (muy importantes para el análisis de regresión lineal):

- $Cov(\hat{\beta}) = s^2(\mathbf{X}^T\mathbf{X})^{-1}$
- Si $E(\epsilon) = \mathbf{0}$, $Cov(\epsilon) = \sigma^2 I$ y $E(\epsilon_i^4) = 3\sigma^4$ para $\mathbf{y} = \mathbf{X}\beta + \epsilon$, entonces s^2 es el mejor estimador cuadrático no sesgado de σ^2

3.2. Regresión lineal y predicción de energía eléctrica generada

Para el diseño del predictor se empleó el siguiente dataset `https://www.kaggle.com/jorgesandoval/wind-power-generation`, este dataset proporciona la potencia real [TVA] producida cada 15 min por una central de energía eólica. De esta forma:

- Variable observada x: tiempo en minutos (cada 15 min).
- Variable predecida Y: potencia acumulada

Para la obtención de la potencia acumulada Y debe aplicarse la siguiente asignación de valores:

$$Y[i] := y[i] + Y[i-1] \tag{3.12}$$

Con $Y[0] = 0$ y y la potencia dada por el data set en [TVA] (potencia activa).
El objetivo de emplear la potencia acumulada para el análisis de regresión lineal es el de predecir la energía total generada en las siguientes horas, días, semanas o incluso meses, también puede obtenerse la potencia instantánea generada $[TV]$ empleando la siguiente asignación de valores:

$$y_{pred}[i] := Y_{pred}[i] - Y_{pred}[i-1] \tag{3.13}$$

Recordemos que esta potencia es clave para satisfacer la energía mínima requerida por los clientes.

CAPÍTULO

4

RESULTADOS DEL PREDICTOR EN PYTHON

Para poder ver el código empleado, puede accederse al siguiente cuaderno de Colab `https://colab.research.google.com/drive/11acqOYjtuSixn4IDXtLJw9QDti1PNwmB?usp=sharing` (cargar el API de su cuenta Kaggle).

4.1. Visualización de los resultados

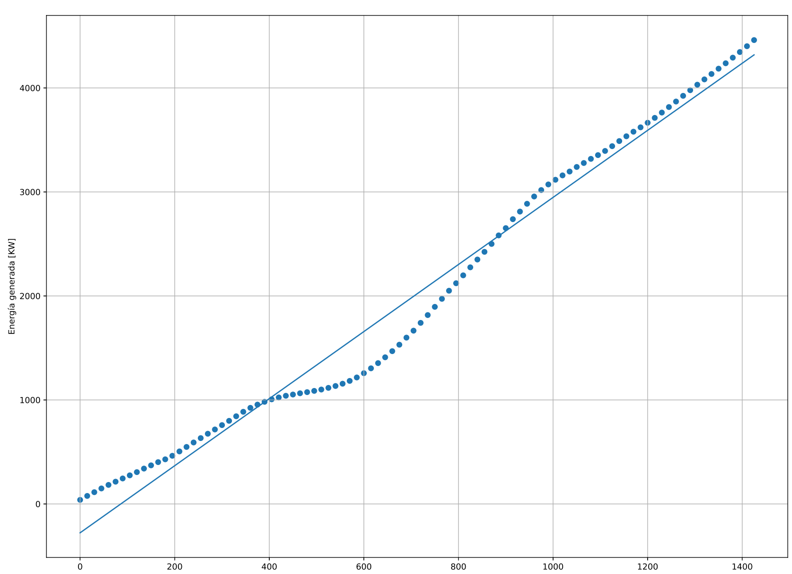


Figura 4.1: Predicción-entrenamiento-27/08/2019-Score= 0,98

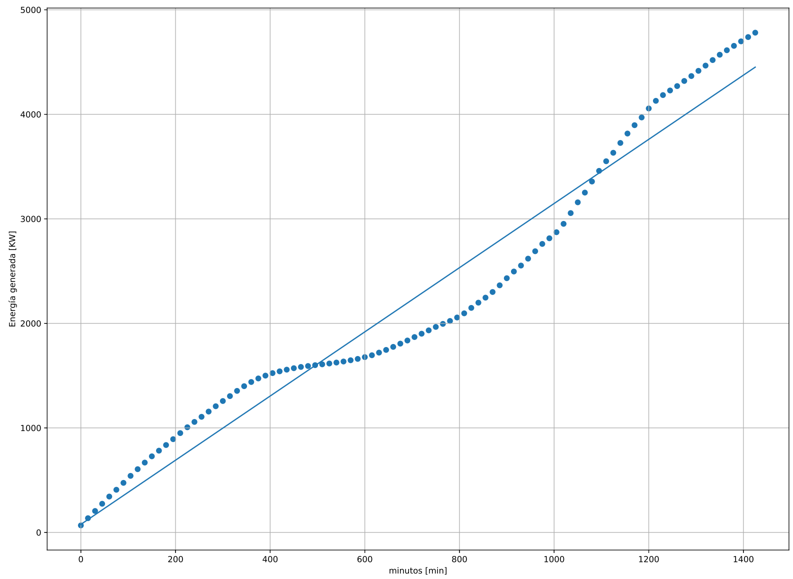


Figura 4.2: Predicción-test-28/08/2019-Score= 0,92

CAPÍTULO

5

CONCLUSIONES

5.1. Conclusiones

- El predictor lineal diseñado tiene un score de 0.92 para la generación de energía eléctrica del día siguiente, por lo cual es recomendable emplear este predictor para predecir la energía total generada en un sistema de molinos.

- La teoría presentada en este trabajo también puede ser aplicada para diseñar un predictor de generación de energía eléctrica de cualquier tipo.

- Pese a la dinámica compleja del molino (sistema electromecánico), se ha visto que la técnica de regresión lineal es eficiente para predecir la dinámica de estos sistemas.

5.2. Recomendaciones

- El dataset empleado solo cuenta con una variable de regresión, por tanto, si este proyecto quisiera aplicarse a las industrias de generación de energía eléctrica de la sociedad Boliviana (hidraúlica, eólica, solar) debería considerarse agregar más variables regresoras, pues la teoría presentada en este se extiende a n regresores.

- Si bien la regresión lineal puede aplicarse de forma eficiente para predecir la energía total generada, no se tiene la certidumbre de que la predicción de la potencia instantánea sea predecida de forma óptima, por tal motivo se recomienda realizar una comparación de score entre la técnica de regresión lineal mostrada en este trabajo con una de predicción basada en forecasting (altamente usada en este ámbito).

- El pre procesamiento de los datos resultó fundamental para obtener el resultado presentado en este trabajo (energía acumulada), de esta forma se recomienda pre procesar los datos para tener un comportamiento lineal entre las variables regresoras y la variable a predecir.

BIBLIOGRAFÍA

[1] Fundamentos de Circuitos Electricos (5a. Ed.).,Alexander, C.K.S., https://books.google.com.bo/books?id=hsaWnQAACAAJ, 2013, McGraw-Hill Interamericana

[2] Matrix Analysis for Statistics, Schott, J.R., Wiley Series in Probability and Statistics, https://books.google.com.bo/books?id=Y2PpCgAAQBAJ, 2016, Wiley.