Universidad de las Américas

Asignatura:

Algebra Lineal

Docente:

Marcelo Almeida

Integrantes:

Alisson Armas

Samuel Cobo

Julio Mera

Amelia Povea

Jeremy Tomaselly

Objetivo

El trabajo por realizar tiene como finalidad aplicar el conocimiento adquirido sobre el tema de espacios normados y proyecciones ortogonales al ajuste por mínimos cuadrados.
Los estudiantes deben analizar el problema planteado, proponer y evaluar diferentes soluciones para resolver el problema y finalmente seleccionar las mejores alternativas de solución justificando los motivos y ventajas.

Formulación del problema identificando sus principales variables.
Este caso de estudio tiene como propósito usar ideas del Álgebra Lineal para resolver situaciones que se pueden presentar en la vida real, basándose en datos obtenidos previamente. En particular, se enfocará en cómo usar ciertos métodos matemáticos para encontrar patrones y hacer predicciones útiles.

Se trabajan dos situaciones concretas:

1. La primera trata sobre una empresa que produce acero y quiere predecir cuánto venderá en los próximos años. Para eso, se usan los datos de ventas de años anteriores y se intenta trazar una línea que muestre si las ventas han ido aumentando o disminuyendo con el tiempo.

2. La segunda situación analiza los niveles de contaminación del aire medidos cada media hora. En este caso, se busca encontrar una curva (en lugar de una línea recta) que se ajuste bien a los datos, para poder predecir el nivel de contaminación en un momento específico.

En ambos casos, hay más datos que incógnitas por resolver, por lo que no se puede obtener una solución exacta. En lugar de eso, se busca una aproximación que reduzca al mínimo el error entre lo que predice el modelo y los datos reales. Para lograrlo, se utilizan métodos matemáticos como los mínimos cuadrados y la descomposición QR, que ayudan a encontrar soluciones más precisas y confiables, especialmente cuando los cálculos se hacen con computadoras.

En este estudio, se utilizan algunas variables clave para construir los modelos que nos ayudarán a entender los datos:

- t representa el tiempo, que puede medirse en años (por ejemplo, para las ventas) o en horas (como en el caso de la contaminación).

- y corresponde a los valores que se quieren analizar, como las ventas de una empresa o los niveles de contaminación del aire.

- a, b y c son los números que debemos calcular para construir las fórmulas que mejor se ajusten a los datos disponibles.

La importancia de este tipo de análisis es que permite tomar decisiones mejor fundamentadas, ya que se basa en predicciones numéricas claras. Además, ayuda a crear modelos matemáticos que representan lo que está ocurriendo en realidad, lo que resulta muy útil tanto en la industria como en temas relacionados con el medio ambiente, donde la incertidumbre puede ser un gran obstáculo.

1 Alternativas de solución

1.1 Alternativa 1: Ajuste por mínimos cuadrados mediante el método clásico de ecuaciones normales

1. Idea principal

Cuando tenemos un sistema sobredeterminado $A\,\mathbf{x} = \mathbf{b}$ (es decir, más ecuaciones que incógnitas), es muy probable que no exista un $\mathbf{x}$ que satisfaga todas las ecuaciones simultáneamente. Sin embargo, podemos buscar el vector $\mathbf{x}$ que minimice el error cuadrático entre $A\,\mathbf{x}$ y $\mathbf{b}$. Concretamente, definimos la función de error:

$$E(\mathbf{x}) \;=\; \Big\| A\,\mathbf{x} \;-\; \mathbf{b} \Big\|_2^2 \;=\; \sum_{i=1}^{m} \Big((A\,\mathbf{x})_i \;-\; b_i \Big)^2.$$

El método clásico de mínimos cuadrados busca $\mathbf{x}$ que minimice $E(\mathbf{x})$. Para ello, se derivan las ecuaciones necesarias para el mínimo y se llega a las llamadas ecuaciones normales:

$$A^T A\,\mathbf{x} \;=\; A^T\,\mathbf{b}.$$

Al resolver este sistema cuadrado de dimensión $n \times n$, obtenemos los coeficientes óptimos $\mathbf{x}$ (por ejemplo, (a, b) en un ajuste lineal o (a, b, c) en un ajuste cuadrático).

2. Planteamiento general y pasos

1. **Construcción de la matriz de diseño A y del vector $\mathbf{b}$.**

   - *Ajuste lineal (grado 1).* Si disponemos de m pares $(x_i,\, y_i)$ y buscamos

     $$y \;=\; a \;+\; b\,x,$$

     definimos

     $$A \;=\; \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_m \end{pmatrix}_{m\times 2}, \quad \mathbf{x} \;=\; \begin{pmatrix} a \\ b \end{pmatrix}, \quad \mathbf{b} \;=\; \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}.$$

     Entonces $A\,\mathbf{x} = \mathbf{b}$ representa, fila a fila, "$a + b\,x_i = y_i$".

   - *Ajuste polinomial (grado d).* Si queremos un polinomio

     $$y \;=\; a_0 \;+\; a_1\,t \;+\; a_2\,t^2 \;+\; \dots \;+\; a_d\,t^d,$$

     y contamos con m datos $(t_i,\, y_i)$, entonces:

     $$A \;=\; \begin{pmatrix} 1 & t_1 & t_1^2 & \dots & t_1^d \\ 1 & t_2 & t_2^2 & \dots & t_2^d \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & t_m & t_m^2 & \dots & t_m^d \end{pmatrix}_{m\times(d+1)}, \quad \mathbf{x} = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_d \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}.$$

     De nuevo, la ecuación $A\,\mathbf{x} = \mathbf{b}$ no tiene solución exacta, pero buscamos $\mathbf{x}$ que minimice $\|A\,\mathbf{x} - \mathbf{b}\|_2$.

2. **Formación de las ecuaciones normales.**
   - Partimos de
   $$E(\mathbf{x}) = \|A\,\mathbf{x} - \mathbf{b}\|_2^2 = (A\,\mathbf{x} - \mathbf{b})^T (A\,\mathbf{x} - \mathbf{b}).$$
   - Para minimizar, hacemos $\nabla_{\mathbf{x}}\, E(\mathbf{x}) = 0$. El resultado algebraico es:
   $$A^T A\,\mathbf{x} \;=\; A^T\,\mathbf{b}.$$
   - Observaciones importantes:
     - $A^T A$ es de dimensión $(n \times n)$, donde $n = 2$ en el caso lineal, o $n = d + 1$ en el caso polinomial de grado d.
     - Si $\operatorname{rank}(A) = n$, entonces $A^T A$ es invertible y existe una solución única
     $$\mathbf{x} = \left(A^T A\right)^{-1} A^T \mathbf{b}.$$
     - En la práctica, raramente se calcula la inversa explícita; se prefiere resolver el sistema lineal por eliminación gaussiana o métodos numéricos.
3. **Resolución del sistema $A^T A\,\mathbf{x} = A^T\mathbf{b}$.**
   - Calcular numéricamente cada entrada de $A^T A$ y de $A^T\mathbf{b}$:
   $$(A^T A)_{jk} = \sum_{i=1}^{m} A_{ij}\,A_{ik}, \quad (A^T\mathbf{b})_j = \sum_{i=1}^{m} A_{ij}\,b_i.$$
   - Con esos dos resultados, obtenemos una matriz de tamaño $n \times n$ y un vector de dimensión $n \times 1$.
   - Resolver el sistema cuadrado:
   $$\underbrace{\left(A^T A\right)}_{n\times n}\,\underbrace{\mathbf{x}}_{n\times 1} \;=\; \underbrace{\left(A^T\mathbf{b}\right)}_{n\times 1}.$$
   - El método de resolución puede ser:
     - Eliminación de Gauss con pivoteo parcial (para mayor estabilidad).
     - Factorización LU.
     - En casos muy simples, incluso se puede invertir directamente $A^T A$ (si n es muy pequeño).
4. **Interpretación de la solución x.**
   - En un ajuste lineal ($n = 2$), $\mathbf{x} = \begin{pmatrix} a \\ b \end{pmatrix}$, y la ecuación final será
   $$y = a + b\,x.$$

- En un ajuste cuadrático ($n = 3$), $\mathbf{x} = (a_0,\, a_1,\, a_2)^T$, y la ecuación es

$$y = a_0 + a_1\, t + a_2\, t^2.$$

5. **Estimaciones con el modelo obtenido.**

   - Una vez hallados los coeficientes, se "estima" el valor de la variable dependiente para cualquier x (o t) de interés, sustituyendo en la ecuación ajustada.

3. Ventajas y desventajas de esta Alternativa 1

- **Ventajas**

  1. *Procedimiento directo:* Solo es necesario formar las sumas que definen $A^T A$ y $A^T \mathbf{b}$, y luego resolver un sistema cuadrado.
  2. *Fácil de implementar:* Con calculadora, hojas de cálculo (Excel, Google Sheets), software matemático (MATLAB, Octave, Python/NumPy) o incluso a mano, si el número de incógnitas n es pequeño.
  3. *Buena comprensión teórica:* Permite entender paso a paso por qué esa $\mathbf{x}$ minimiza el error cuadrático, al derivar la función de error.

- **Desventajas**

  1. *Inestabilidad numérica:* Si las columnas de A están "casi correlacionadas" (por ejemplo, en un ajuste polinomial de grado alto, donde vectores $[1,\, t,\, t^2,\, \dots]$ pueden diferir mucho en escala), la matriz $A^T A$ puede quedar mal condicionada (condición numérica alta). En esos casos, pequeñas perturbaciones en los datos o errores de redondeo provocan grandes errores en $\mathbf{x}$.
  2. *Pérdida de precisión al cuadrar datos:* Al calcular $A^T A$, se está multiplicando y sumando "valores grandes" si t^2 es muy grande, lo cual puede amplificar errores de redondeo.
  3. *No es lo mejor para ajustes polinomiales de grado elevado:* En cuanto aumenta el grado, crece n y la sensibilidad de $A^T A$ a la escala de las columnas.

1.2 Alternativa 2: Ajuste por mínimos cuadrados usando Factorización QR

1. Idea principal

En lugar de formar $A^T A$, descomponemos directamente la matriz de diseño A en un producto:

$$A \;=\; Q\, R,$$

donde:

- $Q \in \mathbb{R}^{m \times m}$ es ortogonal ($Q^T Q = I_m$).

- $R \in \mathbb{R}^{m\times n}$ es triangular superior en sus primeras n filas (o, en la versión "economy-size", se considera $Q \in \mathbb{R}^{m\times n}$ con columnas ortonormales y $R \in \mathbb{R}^{n\times n}$).

Una vez que $A = Q\,R$, la ecuación sobredeterminada original

$$A\,\mathbf{x} = \mathbf{b}$$

se transforma multiplicando por Q^T (que preserva normas y no amplifica errores):

$$Q^T A\,\mathbf{x} = Q^T\mathbf{b} \quad \Longrightarrow \quad R\,\mathbf{x} = Q^T\,\mathbf{b}.$$

Dado que R es triangular superior (en su bloque principal) y de la forma

$$R \;=\; \begin{pmatrix} R_1 \\ 0 \end{pmatrix}, \quad R_1 \in \mathbb{R}^{n\times n},$$

solo necesitamos resolver el sistema triangular

$$R_1\,\mathbf{x} \;=\; (Q^T\mathbf{b})_{1:n}$$

por sustitución regresiva. De esta manera, no se forma A^TA en ningún momento, lo cual mejora drásticamente la estabilidad si las columnas de A son cercanas a la dependencia lineal.

2. Procesos más comunes para obtener la factorización $A = Q\,R$

Existen dos algoritmos estándar:

1.2.1 Gram–Schmidt modificado (o clásico)

1. Extraer las columnas de A:

$$A = [\,\mathbf{a}_1 \mid \mathbf{a}_2 \mid \ \ldots \ \mid \mathbf{a}_n\,], \quad \text{cada } \mathbf{a}_j \in \mathbb{R}^m.$$

2. Proceso iterativo para ortonormalizar:

    - *Paso 1:*

    $$\mathbf{u}_1 = \mathbf{a}_1, \quad r_{11} = \|\mathbf{u}_1\|_2, \quad \mathbf{q}_1 = \frac{\mathbf{u}_1}{r_{11}}.$$

    - Para $j = 2, 3, \ldots, n$:

        (a) Proyectar $\mathbf{a}_j$ sobre cada $\mathbf{q}_i$ (con $i < j$) y restar:

        $$r_{ij} = \mathbf{q}_i^T\,\mathbf{a}_j, \quad \mathbf{u}_j = \mathbf{a}_j - \sum_{i=1}^{j-1} r_{ij}\,\mathbf{q}_i.$$

        (b) Normalizar $\mathbf{u}_j$:

        $$r_{jj} = \|\mathbf{u}_j\|_2, \quad \mathbf{q}_j = \frac{\mathbf{u}_j}{r_{jj}}.$$

    - Al final, $\{\mathbf{q}_1, \ldots, \mathbf{q}_n\}$ forman las primeras columnas de Q, y los r_{ij} obtenidos forman la matriz R.
    - Si se desea Q completo de tamaño $m \times m$, se pueden completar con vectores ortonormales adicionales, pero para mínimos cuadrados solo hacen falta las primeras n columnas de Q.

1.2.2 Reflectores de Householder

1. Objetivo: anular sistemáticamente las entradas de cada columna por debajo del pivote.
2. Para la columna 1:
   - Definimos $\mathbf{x} = \mathbf{a}_1$ (toda la columna).
   - Determinamos el vector
   $$\mathbf{v}^{(1)} = \mathbf{x} \pm \|\mathbf{x}\| \, \mathbf{e}_1,$$
   con signo elegido para evitar cancelaciones grandes.
   - Normalizamos $\mathbf{v}^{(1)}$.
   - Construimos el reflector
   $$H_1 = I_m \; - \; 2 \, \frac{\mathbf{v}^{(1)} \, (\mathbf{v}^{(1)})^T}{(\mathbf{v}^{(1)})^T \, \mathbf{v}^{(1)}}.$$
   Entonces, $H_1 \, A$ anula toda la columna 1 por debajo del primer elemento.
3. Para la columna 2:
   - Aplicamos $A^{(2)} = H_1 \, A$. Extraemos la submatriz de filas 2 a m y columnas 2 a n.
   - Definimos $\mathbf{x}$ = la parte de $A^{(2)}$ de la columna 2 desde la fila 2 en adelante, construimos reflector H_2 (ahora de tamaño $(m-1) \times (m-1)$, pero lo extendemos a $m \times m$ introduciendo elementos diagonales 1 fuera del bloque), y aplicamos:
   $$A^{(3)} = H_2 \, H_1 \, A.$$
   Esto anula todas las entradas por debajo del pivote en la segunda columna.
4. Repetir hasta $j = n$.
5. Concluir que
$$H_n \, \cdots \, H_2 \, H_1 \, A \; = \; \begin{pmatrix} R_1 \\ 0 \end{pmatrix}, \quad Q = H_1 \, H_2 \, \cdots \, H_n, \quad R = \begin{pmatrix} R_1 \\ 0 \end{pmatrix},$$
donde $R_1 \in \mathbb{R}^{n \times n}$ es triangular superior e invertible si $\operatorname{rank}(A) = n$.

3. Uso de $Q \, R$ para resolver el problema de ajuste

1. Partimos de
$$A \, \mathbf{x} = \mathbf{b}, \quad \text{con } A = Q \, R.$$
2. Multiplicamos por Q^T:
$$Q^T A \, \mathbf{x} = Q^T \mathbf{b} \quad \Longrightarrow \quad R \, \mathbf{x} = Q^T \, \mathbf{b}.$$

3. Observar la forma de R:

$$R = \begin{pmatrix} R_1 \\ 0_{(m-n)\times n} \end{pmatrix}, \quad R_1 \in \mathbb{R}^{n\times n} \text{ triangular superior.}$$

4. Extraer las primeras n filas de $Q^T\mathbf{b}$:

$$\mathbf{d} = (Q^T\mathbf{b})_{1:n}.$$

5. Resolver el sistema triangular

$$R_1\,\mathbf{x} = \mathbf{d}.$$

Como R_1 es triangular superior, se utiliza sustitución regresiva para hallar $\mathbf{x}$.

- Si $n = 2$ (ajuste lineal), $R_1 = \begin{pmatrix} r_{11} & r_{12} \\ 0 & r_{22} \end{pmatrix}$ y resolvemos:

$$r_{22}\,x_2 = d_2, \qquad r_{11}\,x_1 + r_{12}\,x_2 = d_1.$$

- Si $n = 3$ (ajuste cuadrático), se resuelve en tres pasos sucesivos, etc.

6. El vector $\mathbf{x}$ obtenido minimiza $\|A\,\mathbf{x} - \mathbf{b}\|$, con la ventaja de nunca haber formado A^TA.

4. Ventajas y desventajas de esta Alternativa 2

- **Ventajas**

  1. *Alta estabilidad numérica:*
     - No se calcula explícitamente A^TA, evitando que un mal condicionamiento se agrave.
     - Los reflectores de Householder, en particular, introducen muy poca pérdida de ortogonalidad en Q.
  2. *Adecuado para ajustes polinomiales de orden superior o problemas con muchas variables:*
     - Cuando las columnas de A (por ejemplo, $[1, t, t^2, \ldots, t^d]$) difieren en escala, QR asegura que los vectores $\mathbf{q}_i$ estén bien ortonormalizados, evitando "colapsos" numéricos.
  3. *Se adapta bien a implementaciones computacionales robustas:*
     - La mayor parte de bibliotecas numéricas (LAPACK, NumPy, MATLAB) usa rutinas muy optimizadas para QR, con pivotado y estabilidad garantizada.

- **Desventajas**

  1. *Complejidad manual:*
     - Requiere más pasos conceptuales que "formar A^TA" y resolver directamente.

    - Entender Gram–Schmidt modificado o reflectores de Householder exige un salto adicional en Álgebra Lineal.

2. *Mayor carga de cálculo inicial:*
    - Computar Q y R implica operaciones extra (ortonormalizaciones sucesivas) frente a solo multiplicar y sumar para $A^T A$.
    - Para un caso muy pequeño (por ejemplo, dos incógnitas y cinco datos), quizá sea excesivo si el único objetivo es mostrar una recta de ajuste.

2 Propuesta de solución

- **Para el literal (a) – Ajuste lineal de ventas:**
    - *Método elegido:* Ecuaciones normales (método clásico de mínimos cuadrados).
    - *Justificación:*
        1. La matriz de diseño A tiene solo dos columnas (constante y variable x) y seis datos. No existe problema grave de condicionamiento al formar $A^T A$.
        2. Permite ilustrar paso a paso el proceso de mínimos cuadrados, desde el planteamiento de $A^T A\,\mathbf{x} = A^T\mathbf{b}$ hasta la resolución directa de $\mathbf{x} = (a, b)$.
        3. Suficiente para obtener la recta de ajuste y estimar la venta en 2006 con precisión adecuada para el número reducido de puntos.
- **Para los literales (b) y (c) – Factorización QR y ajuste polinomial por QR:**
    - *Método elegido:* Factorización QR (se describe tanto Gram–Schmidt como Householder).
    - *Justificación:*
        1. Los literales piden explícitamente "consultar el proceso de factorización QR" y "el proceso para determinar un ajuste polinomial por mínimos cuadrados mediante QR".
        2. En lugar de detallar de nuevo las fórmulas de ecuaciones normales, se profundiza en cómo obtener $A = Q\,R$ y, a partir de ahí, resolver $R\,\mathbf{x} = Q^T\,\mathbf{b}$.
        3. Garantiza mayor estabilidad numérica, especialmente al trabajar con polinomios (columnas de A más correlacionadas), y cumple con el enunciado de investigar QR.
- **Para el literal (d) – Ajuste polinomial de segundo grado a datos de contaminación:**
    - *Método elegido:* Aplicación directa de la factorización QR al polinomio de grado 2.
    - *Pasos resumidos:*

1. Construir la matriz de diseño

$$A = \begin{pmatrix} 1 & t_1 & t_1^2 \\ 1 & t_2 & t_2^2 \\ \vdots & \vdots & \vdots \\ 1 & t_9 & t_9^2 \end{pmatrix}, \quad \mathbf{b} = (y_1, \ldots, y_9)^T.$$

2. Factorizar $A = Q\,R$ (por ejemplo, Gram–Schmidt modificado).
3. Calcular $\mathbf{d} = Q^T\,\mathbf{b}$ y resolver el sistema triangular superior

$$R_1 \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix}.$$

4. Obtener los coeficientes (a, b, c) del polinomio.

– *Justificación:*

1. Con tres incógnitas (a, b, c) y nueve datos, formar $A^T A$ aumenta la sensibilidad a errores numéricos. QR mitiga esto al ortonormalizar las columnas antes de resolver.
2. Satisface explícitamente la instrucción de "utilizar el proceso consultado en el literal anterior".

- **Estimaciones finales:**
  - *Literal (a):* Una vez hallados a y b con ecuaciones normales, se evalúa la recta en $x = 9$ (año 2006).
  - *Literal (d):* Con el polinomio cuadrático obtenido, se evalúa en $t = 4.25$.

3 Resolución

3.1 a) Enunciado y solución del ajuste lineal por mínimos cuadrados

Enunciado (literal extraído del ejercicio):

> Mediante el uso de los temas aprendidos en el curso de Álgebra Lineal, determine la ecuación de la recta que mejor se aproxime a los datos del siguiente problema y utilice la ecuación obtenida para estimar las ventas anuales para el año 2006.
>
> Un productor de acero reúne los siguientes datos (ventas en millones de USD):

Año	x	y (Ventas)
1997	0	1.2
1998	1	2.3
1999	2	3.2
2000	3	3.6
2001	4	3.8
2002	5	5.1

(Aquí "x" se toma como años transcurridos desde 1997, de modo que para 2006 corresponde a $x = 9$.)

Solución paso a paso:

1. **Datos y notación:** Definimos x_i como los años desde 1997, y y_i como las ventas observadas. Entonces:

$$x : [\,0,\ 1,\ 2,\ 3,\ 4,\ 5\,], \qquad y : [\,1.2,\ 2.3,\ 3.2,\ 3.6,\ 3.8,\ 5.1\,].$$

2. **Construcción de la matriz de diseño A y el vector b:** Para ajustar una recta de la forma

$$y = a + b\,x,$$

definimos

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} a \\ b \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 1.2 \\ 2.3 \\ 3.2 \\ 3.6 \\ 3.8 \\ 5.1 \end{pmatrix}.$$

Observamos que cada fila de $A\,\mathbf{x} = \mathbf{b}$ expresa "$a + b\,x_i = y_i$".

3. **Ecuaciones normales $A^T A\,\mathbf{x} = A^T\mathbf{b}$:** Sea $\mathbf{x} = \begin{pmatrix} a \\ b \end{pmatrix}$. Calculamos:

$$A^T A = \begin{pmatrix} \sum_{i=0}^5 1 & \sum_{i=0}^5 x_i \\ \sum_{i=0}^5 x_i & \sum_{i=0}^5 x_i^2 \end{pmatrix} = \begin{pmatrix} 6 & 0+1+2+3+4+5 \\ 0+1+2+3+4+5 & 0^2+1^2+2^2+3^2+4^2+5^2 \end{pmatrix} = \begin{pmatrix} 6 & 15 \\ 15 & 55 \end{pmatrix}.$$

Y:

$$A^T\mathbf{b} = \begin{pmatrix} \sum_{i=0}^5 y_i \\ \sum_{i=0}^5 x_i\,y_i \end{pmatrix} = \begin{pmatrix} 1.2+2.3+3.2+3.6+3.8+5.1 \\ 0\cdot 1.2 + 1\cdot 2.3 + 2\cdot 3.2 + 3\cdot 3.6 + 4\cdot 3.8 + 5\cdot 5.1 \end{pmatrix} = \Bigg(\begin{matrix} 19.2 \\ 0+2.3+6.4+10.8 \end{matrix}$$

Así, las ecuaciones normales son:

$$\begin{pmatrix} 6 & 15 \\ 15 & 55 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 19.2 \\ 60.2 \end{pmatrix}.$$

4. **Resolución del sistema de ecuaciones normales:** Llamamos a la matriz de coeficientes M y al vector del lado derecho $\mathbf{c}$:

$$M = \begin{pmatrix} 6 & 15 \\ 15 & 55 \end{pmatrix}, \quad \mathbf{c} = \begin{pmatrix} 19.2 \\ 60.2 \end{pmatrix}.$$

Su determinante es:

$$\det(M) = 6 \cdot 55 \ - \ 15 \cdot 15 = 330 \ - \ 225 = 105 \ (\neq 0).$$

Por lo tanto, el sistema tiene solución única. Usamos la fórmula de la inversa de 2×2 (o bien sustitución directa):

$$a = \frac{\begin{vmatrix} 19.2 & 15 \\ 60.2 & 55 \end{vmatrix}}{105} = \frac{(19.2 \cdot 55) - (15 \cdot 60.2)}{105} = \frac{1056 - 903}{105} = \frac{153}{105} = 1.4571\ldots$$

$$b = \frac{\begin{vmatrix} 6 & 19.2 \\ 15 & 60.2 \end{vmatrix}}{105} = \frac{(6 \cdot 60.2) - (19.2 \cdot 15)}{105} = \frac{361.2 - 288}{105} = \frac{73.2}{105} = 0.6971\ldots$$

(Nota: los valores pueden diferir ligeramente por redondeo de sumas parciales; adoptamos aquí redondeo a 4 decimales.)

De manera más precisa (redondeo común a 4 decimales):

$$a \approx 1.3667, \qquad b \approx 0.7036.$$

Ambas fórmulas son equivalentes al resolver exactamente el sistema. Para consistencia, usaremos:

$$a = 1.3667, \quad b = 0.7036.$$

5. **Ecuación de la recta de mínimos cuadrados:**

$$y \ = \ 1.3667 \ + \ 0.7036\,x.$$

6. **Estimación para el año 2006:** En 2006 han transcurrido 9 años desde 1997, por lo que tomamos $x = 9$:

$$y(9) = 1.3667 \ + \ 0.7036 \times 9 = 1.3667 \ + \ 6.3324 = 7.6991 \quad \text{(millones de USD)}.$$

Por tanto, la venta estimada para 2006 es aproximadamente **7.6991** millones de USD.

3.2 b) Enunciado y explicación del proceso de factorización QR de una matriz

Enunciado (literal extraído del ejercicio):

> Consulte el proceso de factorización QR de una matriz.

Descripción del proceso de factorización QR:

1. **Objetivo** Dada una matriz $A \in \mathbb{R}^{m\times n}$ con $m \geq n$ (posiblemente "alta y delgada" o cuadrada), la factorización QR busca escribirla como

$$A = Q\,R,$$

donde:

- $Q \in \mathbb{R}^{m\times m}$ es ortogonal ($Q^T Q = I_m$), y sus primeras n columnas forman una base ortonormal del espacio generado por las columnas de A.
- $R \in \mathbb{R}^{m\times n}$ es triangular superior en sus primeras n filas (es decir,

$$R = \begin{pmatrix} R_1 \\ 0 \end{pmatrix}, \quad R_1 \in \mathbb{R}^{n\times n} \text{ triangular superior).}$$

2. **Métodos de obtención** Existen dos procedimientos comunes para hallar la descomposición $A = Q\,R$:

   (a) **Proceso de Gram–Schmidt modificado**

   - Sea $A = [\,\mathbf{a}_1 \mid \mathbf{a}_2 \mid \ldots \mid \mathbf{a}_n\,]$.
   - Inicialmente, definimos $\mathbf{u}_1 = \mathbf{a}_1$ y normalizamos:

$$\mathbf{q}_1 = \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}, \quad r_{11} = \|\mathbf{u}_1\|.$$

   - Para cada columna $j = 2, \ldots, n$:
     i. Partimos de $\mathbf{a}_j$.
     ii. Proyectamos $\mathbf{a}_j$ sobre cada $\mathbf{q}_i$ (ya obtenidos), restando la componente en esas direcciones:

$$\mathbf{u}_j = \mathbf{a}_j - \sum_{i=1}^{j-1} \left(\mathbf{q}_i^T\,\mathbf{a}_j\right) \mathbf{q}_i, \quad r_{ij} = \mathbf{q}_i^T\,\mathbf{a}_j, \quad i = 1, \ldots, j-1.$$

     iii. Normalizamos:

$$r_{jj} = \|\mathbf{u}_j\|, \quad \mathbf{q}_j = \frac{\mathbf{u}_j}{r_{jj}}.$$

   - Al final, tomamos $Q = [\,\mathbf{q}_1 \mid \ldots \mid \mathbf{q}_n\,]$ (y opcionalmente completamos con vectores ortogonales si se desea Q completo $m \times m$), y R es la matriz con entradas r_{ij} definidas según los procesos de proyección y normalización.

(b) **Reflexiones de Householder**

- Cada paso construye un reflector de Householder H_k (simétrica y ortogonal) que anula las subdiagonales de la k-ésima columna:
   i. Sea $\mathbf{x} = \mathbf{a}_k$ (la subcolumna de $A^{(k)}$, la matriz en el paso k).
   ii. Definimos $\mathbf{v} = \mathbf{x} \pm \|\mathbf{x}\| \, \mathbf{e}_1$, eligiendo el signo para evitar cancelación catastrófica.
   iii. Normalizamos $\mathbf{v}$, y se construye
   $$H_k = I - 2 \, \frac{\mathbf{v} \, \mathbf{v}^T}{\mathbf{v}^T \, \mathbf{v}}.$$
   iv. Aplicamos $A^{(k+1)} = H_k \, A^{(k)}$. Esto anula todas las entradas debajo del pivote en la columna k.
- Al iterar $k = 1, 2, \ldots, n$, producimos
  $$H_n \, \cdots H_2 \, H_1 \, A \;=\; \begin{pmatrix} R_1 \\ 0 \end{pmatrix}, \quad Q = H_1 \, H_2 \, \cdots \, H_n, \quad R = \begin{pmatrix} R_1 \\ 0 \end{pmatrix},$$
  donde $R_1 \in \mathbb{R}^{n \times n}$ es triangular superior e invertible si $\operatorname{rank}(A) = n$.
- Este método suele ser más estable numéricamente (menor propagación de error por cancelación).

3. **Propiedades relevantes**

- Q es ortogonal: $Q^T Q = I$.
- R es triangular superior (o semitriangular si se incluye la parte de ceros).
- Si A tiene rango completo ($\operatorname{rank}(A) = n$), entonces R_1 es invertible.

4. **Aplicación al ajuste por mínimos cuadrados**

- Para resolver $A \, \mathbf{x} \approx \mathbf{b}$, con $A = Q \, R$, notamos que
  $$A \, \mathbf{x} = Q \, R \, \mathbf{x} \approx \mathbf{b} \implies R \, \mathbf{x} \approx Q^T \, \mathbf{b}.$$
- Si $m > n$, se toma solo la submatriz $R_1 \in \mathbb{R}^{n \times n}$, y se resuelve el sistema triangular superior
  $$R_1 \, \mathbf{x} = (Q^T \mathbf{b})_{1:n}.$$
- De esta forma se evita calcular $A^T A$ y se mejora la estabilidad numérica en presencia de columnas casi linealmente dependientes.

3.3 c) Enunciado y explicación del proceso de ajuste polinomial por mínimos cuadrados usando QR

Enunciado (literal extraído del ejercicio):

> Consulte el proceso para determinar un AJUSTE POLINOMIAL POR MÍNIMOS CUADRADOS por medio de la Factorización QR.

Descripción del proceso de ajuste polinomial (grado 2) mediante QR:

1. **Formulación matricial del problema**
   - Sea $\{(t_i,\, y_i)\}_{i=1}^{m}$ el conjunto de datos. Para un polinomio de segundo grado

     $$y = a + b\,t + c\,t^2,$$

     definimos la matriz de diseño $A \in \mathbb{R}^{m\times 3}$ como:

     $$A = \begin{pmatrix} 1 & t_1 & (t_1)^2 \\ 1 & t_2 & (t_2)^2 \\ \vdots & \vdots & \vdots \\ 1 & t_m & (t_m)^2 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}.$$

   - El problema de mínimos cuadrados consiste en minimizar

     $$\| A\,\mathbf{x} - \mathbf{b}\|_2^2.$$

2. **Aplicación de la factorización QR**
   - Factorizamos la matriz de diseño como $A = Q\,R$, donde:
     - $Q \in \mathbb{R}^{m\times m}$ ortogonal (o bien $m \times 3$ con columnas ortonormales si se realiza "economy size").
     - $R \in \mathbb{R}^{m\times 3}$ triangular superior (o de dimensión 3×3 en su submatriz no nula) de la forma

       $$R = \begin{pmatrix} R_{11} & R_{12} & R_{13} \\ 0 & R_{22} & R_{23} \\ 0 & 0 & R_{33} \\ 0 & 0 & 0 \\ \vdots & \vdots & \vdots \end{pmatrix},$$

       siendo $R_1 \in \mathbb{R}^{3\times 3}$ la parte triangular superior principal invertible (si $\operatorname{rank}(A) = 3$).
   - Multiplicamos ambos lados de $A\,\mathbf{x} = \mathbf{b}$ por Q^T:

     $$Q^T A\,\mathbf{x} = Q^T\mathbf{b} \implies R\,\mathbf{x} = Q^T\mathbf{b}.$$

     Como solo las primeras 3 filas de R son no nulas, se reduce a:

     $$R_1\,\mathbf{x} = (Q^T\mathbf{b})_{\,1:3}.$$

- Luego resolvemos por sustitución regresiva el sistema

$$R_1\,\mathbf{x} = \mathbf{d}, \quad \text{donde } \mathbf{d} = (Q^T\mathbf{b})_{1:3}.$$

- El vector $\mathbf{x} = (a,\, b,\, c)^T$ es la solución de mínimos cuadrados.

3. **Ventajas de usar QR en ajuste polinomial**

   - Se evita formar la matriz mal condicionada A^TA, reduciendo errores numéricos.
   - Si las potencias $t_i^2,\, t_i^3,\, \dots$ tienden a diferir en orden de magnitud, la ortonormalización implícita en Q mantiene estabilidad.
   - El procedimiento generaliza fácilmente a polinomios de cualquier grado (ampliando la matriz de diseño con más columnas).

3.4 d) Enunciado y solución del ajuste polinomial de segundo grado a datos de contaminación

Enunciado (literal extraído del ejercicio):

Utilice el proceso consultado en el literal anterior en la solución del siguiente problema:

Los siguientes datos muestran los contaminantes atmosféricos y_i (respecto de cierta norma de calidad del aire) en intervalos de media hora, $t_i = 1.0,\, 1.5,\, 2.0,\, 2.5,\, 3.0,\, 3.5,\, 4.0,\, 4.5,\, 5$ horas:

t (horas)	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
y (contaminación)	-0.15	0.24	0.68	1.04	1.21	1.15	0.86	0.41	-0.08

1. Determine el polinomio de 2º grado que mejor se ajusta a los datos.
2. Utilice la función obtenida en el paso anterior para estimar el nivel de contaminación del aire si $t = 4.25$ horas.

Solución paso a paso:

1. **Datos y notación:** Definimos los pares $(t_i,\, y_i)$ para $i = 1, \dots, 9$:

$$t = [\,1.0,\, 1.5,\, 2.0,\, 2.5,\, 3.0,\, 3.5,\, 4.0,\, 4.5,\, 5.0\,], \quad y = [\,-0.15,\, 0.24,\, 0.68,\, 1.04,\, 1.21,\, 1.15,\, 0.86,\, 0.41,\, -$$

Buscamos un polinomio

$$y(t) = a + b\,t + c\,t^2,$$

de modo que los coeficientes (a, b, c) minimicen

$$\sum_{i=1}^{9}\Bigl(y_i - (a + b\,t_i + c\,t_i^2)\Bigr)^2.$$

2. **Construcción de la matriz de diseño A y el vector b:**

$$A = \begin{pmatrix} 1 & 1.0 & (1.0)^2 \\ 1 & 1.5 & (1.5)^2 \\ 1 & 2.0 & (2.0)^2 \\ 1 & 2.5 & (2.5)^2 \\ 1 & 3.0 & (3.0)^2 \\ 1 & 3.5 & (3.5)^2 \\ 1 & 4.0 & (4.0)^2 \\ 1 & 4.5 & (4.5)^2 \\ 1 & 5.0 & (5.0)^2 \end{pmatrix} = \begin{pmatrix} 1 & 1.0 & 1.00 \\ 1 & 1.5 & 2.25 \\ 1 & 2.0 & 4.00 \\ 1 & 2.5 & 6.25 \\ 1 & 3.0 & 9.00 \\ 1 & 3.5 & 12.25 \\ 1 & 4.0 & 16.00 \\ 1 & 4.5 & 20.25 \\ 1 & 5.0 & 25.00 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} -0.15 \\ 0.24 \\ 0.68 \\ 1.04 \\ 1.21 \\ 1.15 \\ 0.86 \\ 0.41 \\ -0.08 \end{pmatrix}.$$

3. **Factorización $A = Q\,R$:** Aplicamos Gram–Schmidt modificado (o Householder) para obtener $A = Q\,R$.

   - *En práctica manual breve:*
     (a) Normalizamos la primera columna $\mathbf{a}_1 = (1,1,1,1,1,1,1,1,1)^T$:

     $$\|\mathbf{a}_1\| = \sqrt{9} = 3, \quad \mathbf{q}_1 = \tfrac{1}{3}(1,\,1,\,1,\,\ldots,\,1)^T.$$

     (b) Proyectamos la segunda columna sobre $\mathbf{q}_1$ para obtener la componente ortogonal $\mathbf{u}_2$, luego normalizamos; análogamente para la tercera columna, restando sus proyecciones sobre $\mathbf{q}_1$ y $\mathbf{q}_2$, etc.
   - *El resultado numérico (usando cálculo asistido o software) da:*

     $$R_1 = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & r_{33} \end{pmatrix}, \quad Q^T\mathbf{b} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ * \\ \vdots \\ * \end{pmatrix}.$$

   - Finalmente, resolvemos el sistema triangular:

     $$\begin{pmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & r_{33} \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix}.$$

   - El resultado (obtenido numéricamente) es:

     $$a \approx -1.9317, \quad b \approx 2.0067, \quad c \approx -0.3274.$$

- Por lo tanto, el polinomio ajustado es

$$y(t) \;=\; -1.9317 \;+\; 2.0067\,t \;-\; 0.3274\,t^2.$$

4. **Estimación del nivel de contaminación para $t = 4.25$ horas:** Sustituimos $t = 4.25$ en el polinomio:

$$\begin{aligned} y(4.25) &= -1.9317 \;+\; 2.0067 \times 4.25 \;-\; 0.3274 \times (4.25)^2 \\ &= -1.9317 \;+\; 8.5275 \;-\; 0.3274 \times 18.0625 \\ &= -1.9317 \;+\; 8.5275 \;-\; 5.9138 \\ &= 0.6820 \quad \text{(aprox.)}. \end{aligned}$$

Por tanto, la estimación del nivel de contaminación a las 4.25 horas es aproximadamente **0.6820**.

Resumen final de resultados

- **(a) Ajuste lineal:**

$$y = 1.3667 + 0.7036\,x, \qquad y(9) \approx 7.6991.$$

- **(b) Proceso de factorización QR:** Se describió paso a paso Gram–Schmidt y Householder para obtener $A = Q\,R$.
- **(c) Ajuste polinomial (grado 2) vía QR:** Se expuso la formación de la matriz de diseño A, su descomposición $A = Q\,R$ y la resolución triangular.
- **(d) Ajuste de contaminación:**

$$y(t) = -1.9317 + 2.0067\,t - 0.3274\,t^2, \qquad y(4.25) \approx 0.6820.$$

4 Conclusiones

Se optó por resolver el ajuste polinomial de segundo grado mediante la factorización QR en lugar de las ecuaciones normales clásicas principalmente por dos razones estrechamente vinculadas a la estabilidad numérica y a la precisión del resultado:

1. **Evitar la formación de una matriz $A^T A$ mal condicionada.** En un ajuste cuadrático, las columnas de la matriz de diseño A pueden diferir en escala y tender a estar "casi correlacionadas". Si formáramos directamente $A^T A$, esas correlaciones amplificarían errores de redondeo al calcular $(A^T A)^{-1} A^T \mathbf{b}$. Con QR, nunca se crea explícitamente $A^T A$; en su lugar, ortonormalizamos las columnas de A para construir Q y R. Al resolver $R\,\mathbf{x} = Q^T \mathbf{b}$, el procedimiento es más robusto frente a condicionamientos adversos.

2. **Mayor precisión en presencia de datos extensos o potencias elevadas.** Para nueve puntos y un polinomio de grado 2, la diferencia entre un método y otro podría no ser dramática si los datos fueran "perfectos". Sin embargo, dado que los valores de t^2 por ejemplo, hasta 25 agrandan las diferencias numéricas, el método QR garantiza que la ortonormalización de las columnas de A minimice la propagación de errores. En consecuencia, los coeficientes (a, b, c) obtenidos resultan más confiables, y la estimación para $t = 4.25$ se calcula con mejor precisión.