Ajustes por mínimos cuadrados

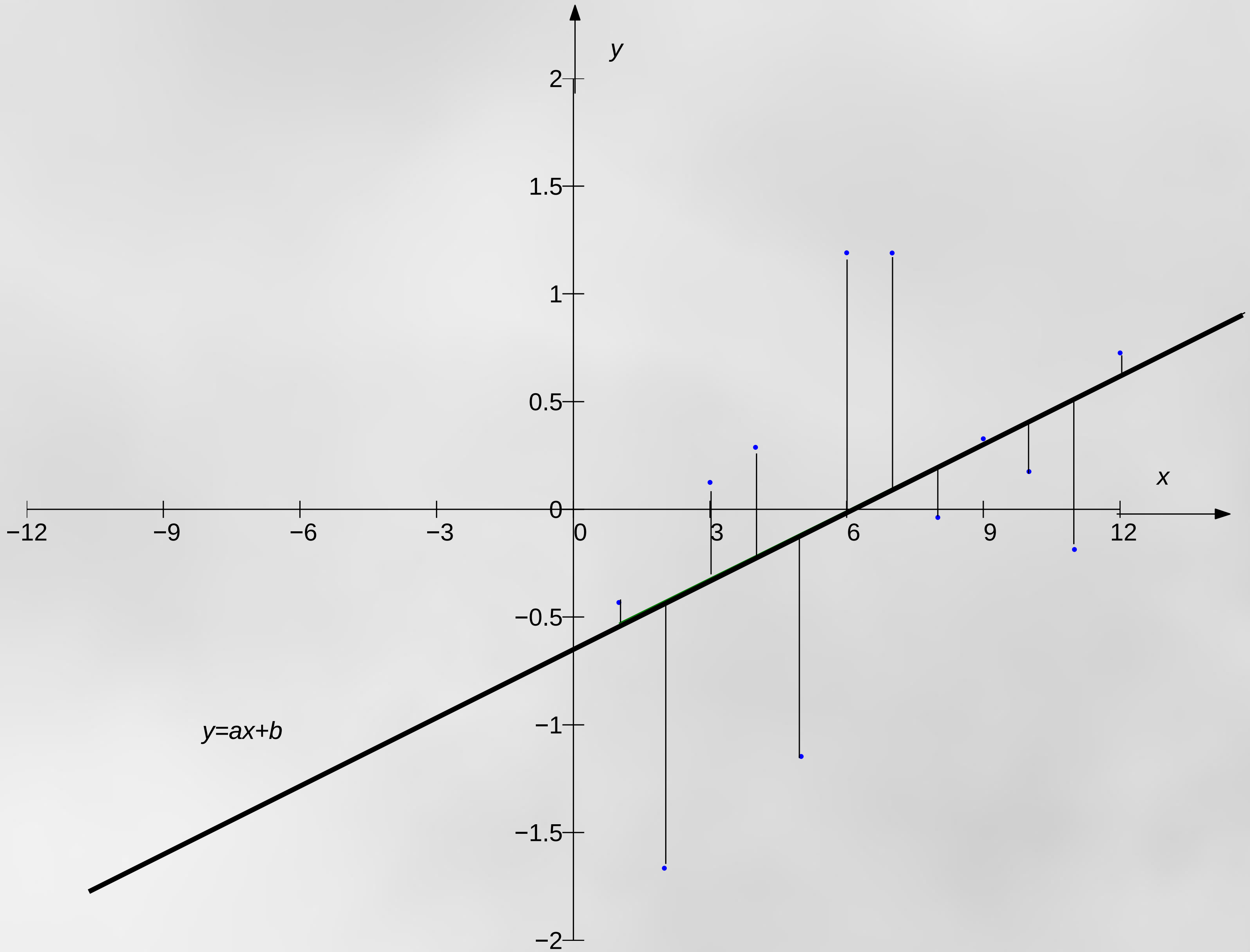
Sean $x_1, x_2, \ldots, x_n$ números reales distintos y sean $y_1, y_2, \ldots, y_n$ números reales.

Ajustes por mínimos cuadrados

Sean $x_1, x_2, \ldots, x_n$ números reales distintos y sean $y_1, y_2, \ldots, y_n$ números reales.

- Problema.- Hallar los coeficientes a y b de la recta $y = ax + b$ que mejor se ajusta a los puntos $(x_1, y_1), \ldots, (x_n, y_n)$. Es decir, buscar el punto (a, b) que hace mínimo el valor de la función

$$E(a,b) := \sum_{i=1}^{n} (ax_i + b - y_i)^2.$$


y
x
2
1.5
1
0.5
0
−0.5
−1
−1.5
−2
−12
−9
−6
−3
0
3
6
9
12
y=ax+b

- Problema.- Hallar el polinomio de segundo grado $y = ax^2 + bx + c$ que mejor se ajusta a los puntos $(x_1, y_1), \ldots, (x_n, y_n)$; en este caso, se debe buscar el mínimo de la función $E(a, b, c)$ de las variables a, b, c, dada por

$$E(a, b, c) := \sum_{i=1}^{n} (ax_i^2 + bx_i + c - y_i)^2.$$

- Problema.- Hallar la elipse $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$ que mejor se ajusta a los puntos $(x_1, y_1), \ldots, (x_n, y_n)$ en el sentido de los mínimos cuadrados.

Ajustes de planos

Dados n puntos en el espacio $(x_1, y_1, z_1), \ldots, (x_n, y_n, z_n)$ hallar el plano $z = ax + by + c$ que minimice el error cuadrático total

$$E(a, b, c) := \sum_{i=1}^{n} (ax_i + by_i + c - z_i)^2.$$

Sistemas lineales incompatibles

Sabemos que un sistema de ecuaciones lineales

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \dotfill \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m, \end{cases}$$

tiene soluciones (o es compatible) si se cumple el Teorema de Rouché-Frobenius.

Este teorema dice que el sistema tiene soluciones si y solo si los rangos de las matrices

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix}$$

son iguales.

Ejemplo 1.- El sistema

$$\begin{cases} 3x_1 + 2x_2 + x_3 = -2, \\ 9x_1 + 6x_2 + 3x_3 = -6, \end{cases}$$

tiene solución,

Ejemplo 1.- El sistema

$$\begin{cases} 3x_1 + 2x_2 + x_3 = -2, \\ 9x_1 + 6x_2 + 3x_3 = -6, \end{cases}$$

tiene solución, pues

$$\operatorname{rg}\begin{pmatrix} 3 & 2 & 1 \\ 9 & 6 & 3 \end{pmatrix} = 1 \text{ y } \operatorname{rg}\begin{pmatrix} 3 & 2 & 1 & -2 \\ 9 & 6 & 3 & -6 \end{pmatrix} = 1.$$

Ejemplo 2.- El sistema

$$\begin{cases} 3x_1 + 2x_2 + x_3 = -2, \\ 9x_1 + 6x_2 + 3x_3 = -6.01, \end{cases}$$

no tiene ninguna solución,

Ejemplo 2.- El sistema

$$\begin{cases} 3x_1 + 2x_2 + x_3 = -2, \\ 9x_1 + 6x_2 + 3x_3 = -6.01, \end{cases}$$

no tiene ninguna solución, pues

$$\mathrm{rg}\begin{pmatrix} 3 & 2 & 1 \\ 9 & 6 & 3 \end{pmatrix} = 1 \text{ y } \mathrm{rg}\begin{pmatrix} 3 & 2 & 1 & -2 \\ 9 & 6 & 3 & -6.01 \end{pmatrix} = 2.$$

Ejemplo 2.- El sistema

$$\begin{cases} 3x_1 + 2x_2 + x_3 = -2, \\ 9x_1 + 6x_2 + 3x_3 = -6.01, \end{cases}$$

no tiene ninguna solución, pues

$$\mathrm{rg}\begin{pmatrix} 3 & 2 & 1 \\ 9 & 6 & 3 \end{pmatrix} = 1 \text{ y } \mathrm{rg}\begin{pmatrix} 3 & 2 & 1 & -2 \\ 9 & 6 & 3 & -6.01 \end{pmatrix} = 2.$$

¿Es satisfactoria esta situación? Cualquiera diría que este segundo sistema “casi” tiene solución.

Esto llegará a ser cierto si en las ecuaciones cambiamos el signo igual $=$ por el signo de aproximadamente igual $\simeq$ (ó $\approx$ en inglés).

$$\begin{cases} 3x_1 + 2x_2 + x_3 \simeq -2, \\ 9x_1 + 6x_2 + 3x_3 \simeq -6.01, \end{cases}$$

Consideremos un caso más general, hallar las soluciones que mejor satisfacen al sistema

$$\begin{cases} a_{11}x_1 + a_{12}x_2 \simeq b_1 \\ a_{21}x_1 + a_{22}x_2 \simeq b_2 \\ a_{31}x_1 + a_{32}x_2 \simeq b_3 \end{cases}$$

Este sistema lo podemos reescribir en forma de vectores columnas así:

$$x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix} \simeq \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

Este sistema lo podemos reescribir en forma de vectores columnas así:

$$x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix} \simeq \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

De esta forma el problema se puede reformular de este modo: Hallar el vector del “plano” (subespacio vectorial de dimensión 2) engendrado por los vectores

$$\begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix} \text{ y } \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix}$$

que esté lo más cerca posible del vector

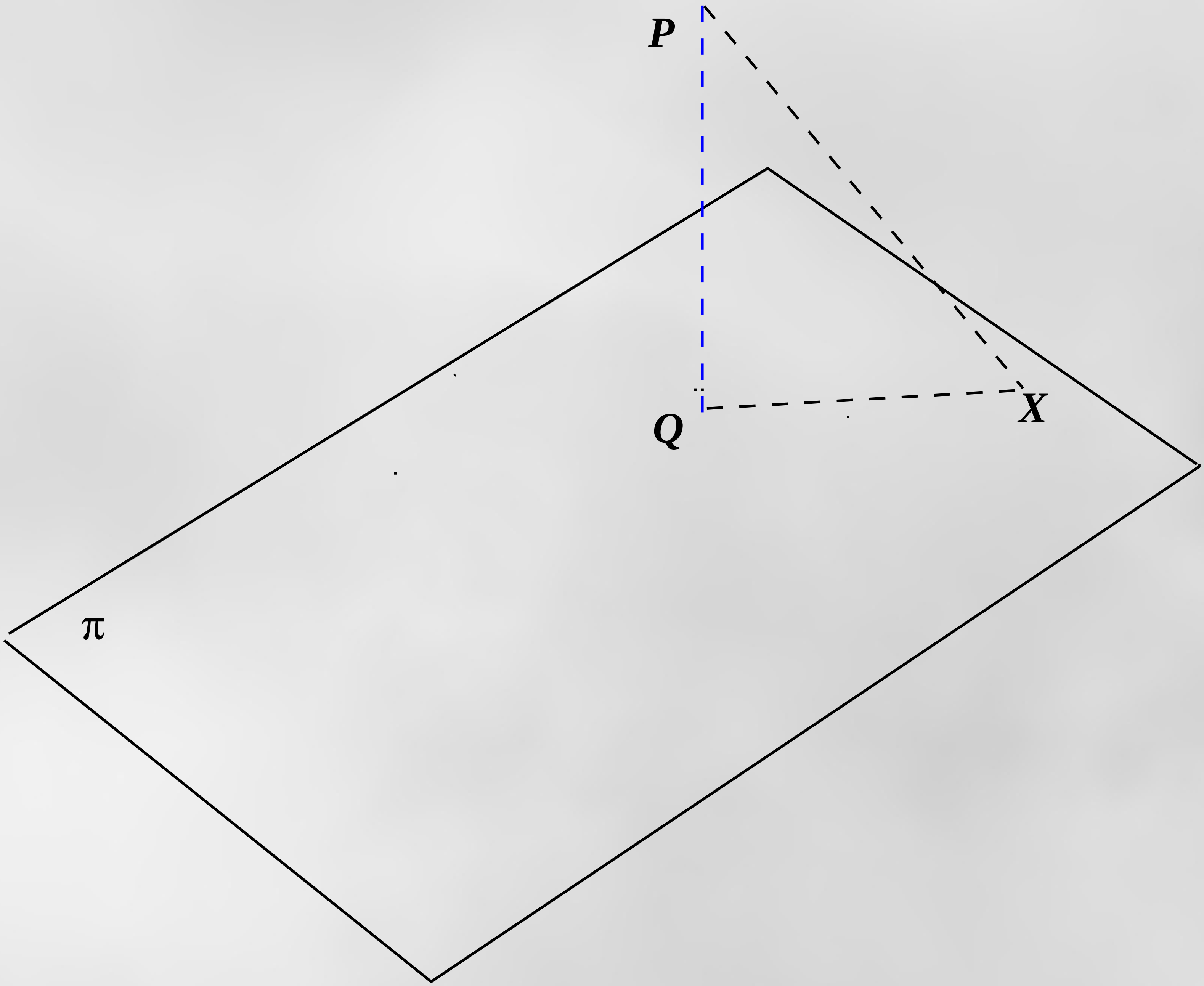
$$\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

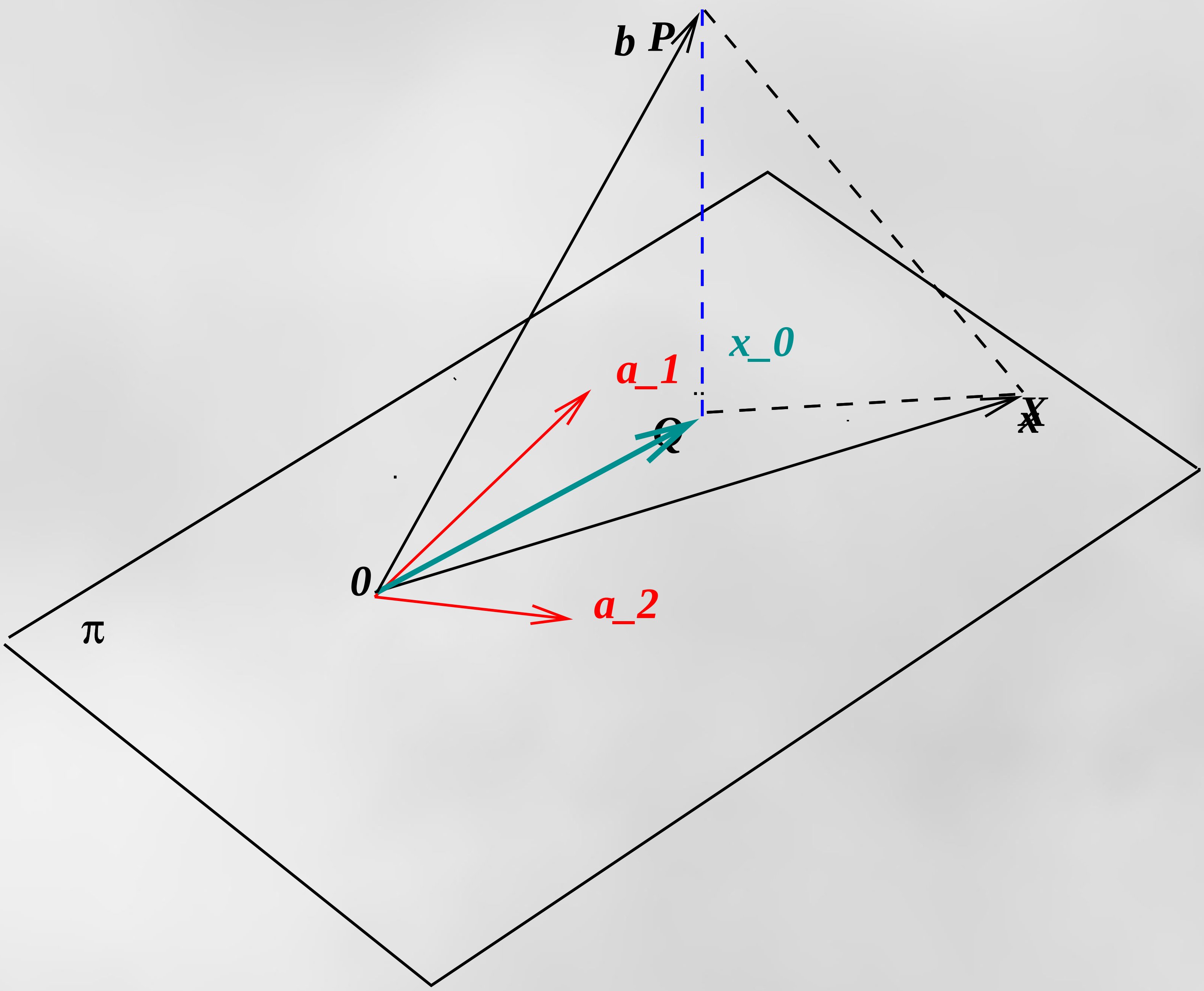
Con lo cual este problema es equivalente al problema de hallar la distancia de un punto a un plano; es decir, dados un punto P y un plano π hallar el punto Q de π que hace mínima la distancia $\overline{PX}$ de P a un punto cualquiera X de π.

Con lo cual este problema es equivalente al problema de hallar la distancia de un punto a un plano; es decir, dados un punto P y un plano π hallar el punto Q de π que hace mínima la distancia $\overline{PX}$ de P a un punto cualquiera X de π.

Es fácil averiguar la solución utilizando el teorema de Pitágoras, pues Q viene dado por la intersección de la recta que pasa por P y es perpendicular a π con este plano. La mínima distancia es $\overline{PQ}$.

Podemos observar dos figuras que ilustran esta idea.


P
Q
X
π


b P
x_0
a_1
Q
X
x
0
a_2
π

Interpretación geométrica del ajuste por mínimos cuadrados

Dados los puntos $(x_1, y_1), \ldots, (x_n, y_n)$ la búsqueda de (a, b) tal que la recta $y = ax + b$ sea la que mejor se ajusta a los puntos en el sentido de los mínimos cuadrados, puede intepretarse de la manera siguiente.

Interpretación geométrica del ajuste por mínimos cuadrados

Dados los puntos $(x_1, y_1), \ldots, (x_n, y_n)$ la búsqueda de (a, b) tal que la recta $y = ax + b$ sea la que mejor se ajusta a los puntos en el sentido de los mínimos cuadrados, puede intepretarse de la manera siguiente.

Si miramos la fórmula

$$E(a, b) := \sum_{i=1}^{n} (ax_i + b - y_i)^2.$$

vemos que se trata de hacer lo más pequeñas posible las siguientes cantidades simultáneamente,

$$(ax_1 + b - y_1)^2$$
$$(ax_2 + b - y_2)^2$$
$$\vdots$$
$$(ax_n + b - y_n)^2$$

Lo cual es equivalente a hallar la combinación lineal

$$a \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + b \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \text{ más próxima al vector } \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$