

Droite de régression (moindres carrés)

A démontrer en classe :

La droite qui *minimise* la somme des carrés des distances *verticales* ($y - y_i$) par rapport à un nuage de points $\{P_i: (x_i; y_i)\}_{i \in \{1 \dots n\}}$

a comme équation : $y = \frac{\text{Covar}(x, y)}{\text{Var}(x)}(x - \bar{x}) + \bar{y}$

avec : $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$ (moyennes) et $\bar{y} = \frac{\sum_{i=1}^n y_i}{n}$

$$\text{Var}(x) = V_{xx} = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$$

$$\text{Var}(y) = V_{yy} = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n}$$

$$\text{Covar}(x, y) = V_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n}$$

De même, la droite qui *minimise* la somme des carrés des distances *horizontales* ($x - x_i$) par rapport au nuage de points $\{P_i: (x_i; y_i)\}_{i \in \{1 \dots n\}}$

a comme équation :

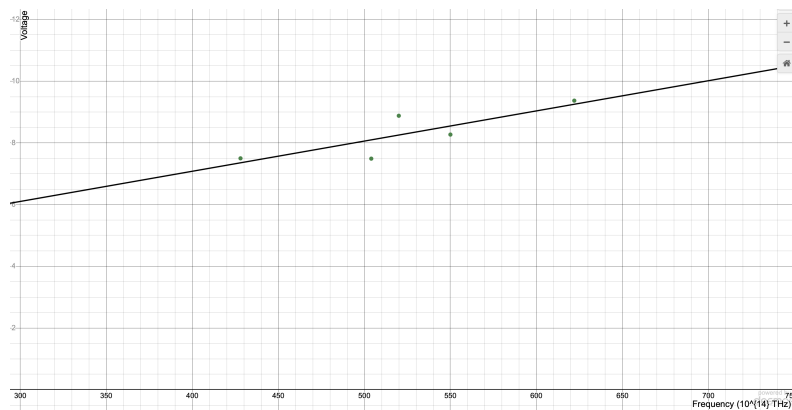
Exemple de calcul

Frequency(f)[THz]	Power(P)[W]	Voltage(V)[V]
622	2.9	9.37
550	2.0	8.27
520	3.1	8.88
504	2.1	7.49
428	1.7	7.50

Mesure de la *tension* (volt) fournie par un panneau solaire pour différentes fréquences d'un spot lumineux (à puissance électrique constante)



- 1) Quelle est l'équation de la droite qui passe *au mieux* parmi les cinq points expérimentaux ?
- 2) Selon le modèle d'une relation *linéaire*:
 - i) Quelle devrait-être la tension électrique fournie par le panneau solaire si la fréquence lumineuse est celle de la raie D₁ du sodium ? ($\lambda_1 = 589,5924 \text{ nm}$)
 - ii) Quelle fréquence donnerait une tension de 10v ?
 - iii) Qu'est-ce qui surprend (physiquement) ?



Réponses

1)

– Première méthode (réponses ici)

i	1	2	3	4	5	Indicateur
x	4.28	5.04	5.20	5.50	6.22	$\bar{x} =$
$x - \bar{x}$						
$(x - \bar{x})^2$						$V_{xx} =$
y	7.5	7.49	8.88	8.27	9.37	$\bar{y} =$
$y - \bar{y}$						
$(y - \bar{y})^2$						(V_{yy})
$(x - \bar{x})(y - \bar{y})$						$V_{xy} =$

– Deuxième méthode (réponses ici)

i	1	2	3	4	5	moyenne	moyenne au carré	moyenne des carrés	Variance
x	4.28	5.04	5.20	5.50	6.22	$\bar{x} =$	$\bar{x}^2 =$	$\bar{x}^2 =$	$V_x =$
y	7.5	7.49	8.88	8.27	9.37	$\bar{y} =$			

– Troisième méthode : Utiliser votre calculatrice (voir [vidéo ici](#))

2) i)

ii)

iii)

Questions complémentaires

- Quelle est la différence entre une *interpolation* et une *extrapolation* ?
- Comment démontrer la relation entre $V_x, \bar{x}^2, \bar{x}^2$ sur laquelle repose la seconde méthode ?
- Qu'est-ce que l'*écart type* d'une variable statistique ? Quelle est son unité ?
- Que vaut la *moyenne géométrique* des droites des pentes des deux droites des moindres carrés ?
- Quel *critère* utiliser pour savoir si le *modèle linéaire* est plus ou moins *plausible* ?