Chemometrie, ISBN: 9789491764332
Vierde herziene druk
www.syntaxmedia.nl

## Uitwerkingen
In de uitgewerkte voorbeelden worden vanwege de leesbaarheid afgeronde tussenresultaten gepresenteerd. De eindresultaten zijn echter altijd berekend zonder tussentijds afronden.

## Hoofdstuk 5

### Antwoord 5.1

Eerst worden voor alle mogelijke combinaties van twee meetpunten de hellingen $\Delta y/\Delta x$ uitgerekend. Daarvan wordt de mediaan bepaald. Deze mediaan is gelijk aan de helling $b_1$ van de robuuste regressielijn die is bepaald volgens de methode van Theil, $b_1$ = 0,0295. Zie onderstaande tabel.

| Combinatie van 2 meetpunten | | x1 | x2 | x2-x1 | y1 | y2 | y2-y1 | helling |
|---|---|---|---|---|---|---|---|---|
| 1 | 2 | 5 | 10 | 5 | 0,16 | 0,29 | 0,13 | 0,0260 |
| 1 | 3 | 5 | 15 | 10 | 0,16 | 0,44 | 0,28 | 0,0280 |
| 1 | 4 | 5 | 20 | 15 | 0,16 | 0,58 | 0,42 | 0,0280 |
| 1 | 5 | 5 | 25 | 20 | 0,16 | 0,83 | 0,67 | 0,0335 |
| 2 | 3 | 10 | 15 | 5 | 0,29 | 0,44 | 0,15 | 0,0300 |
| 2 | 4 | 10 | 20 | 10 | 0,29 | 0,58 | 0,29 | 0,0290 |
| 2 | 5 | 10 | 25 | 15 | 0,29 | 0,83 | 0,54 | 0,0360 |
| 3 | 4 | 15 | 20 | 5 | 0,44 | 0,58 | 0,14 | 0,0280 |
| 3 | 5 | 15 | 25 | 10 | 0,44 | 0,83 | 0,39 | 0,0390 |
| 4 | 5 | 20 | 25 | 5 | 0,58 | 0,83 | 0,25 | 0,0500 |
| | | | | | | | mediaan | 0,0295 |

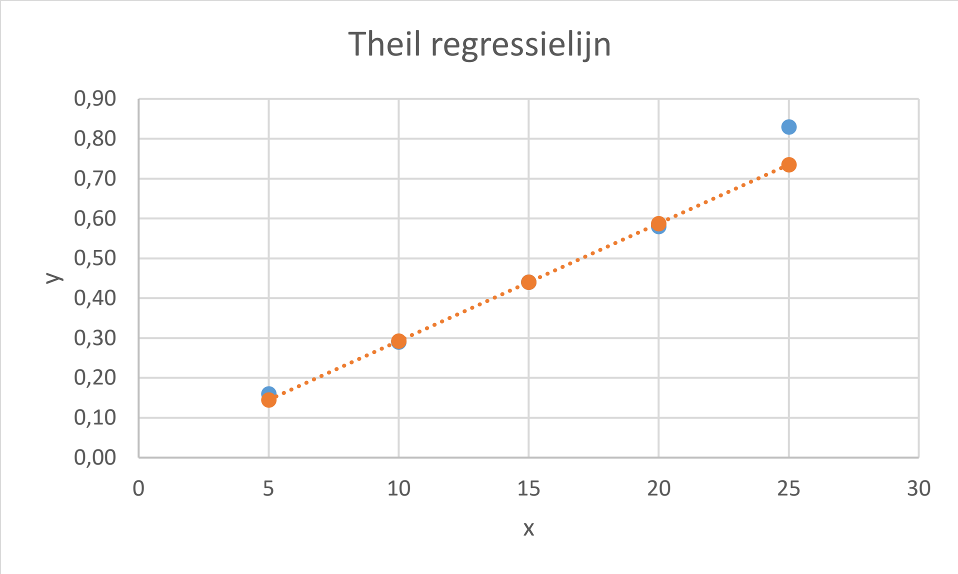
Vervolgens wordt door alle meetpunten een rechte lijn getrokken met een helling $b_1$ = 0,0295. Van elke rechte lijn wordt de as-afsnede berekend met $b_{0,i} = y_i - b_1 x_i$.
Daarna wordt de mediaan bepaald van deze as-afsneden. Deze mediaan is gelijk aan de as-afsnede $b_0$ van de robuuste regressielijn die is bepaald volgens de methode van Theil, $b_0$ = -0,0025. Zie onderstaande tabel.

| nr | x | y | as-afsnede |
|---|---|---|---|
| 1 | 5 | 0,16 | 0,0125 |
| 2 | 10 | 0,29 | -0,0050 |
| 3 | 15 | 0,44 | -0,0025 |
| 4 | 20 | 0,58 | -0,0100 |
| 5 | 25 | 0,83 | 0,0925 |
| | | mediaan | -0,0025 |

De vergelijking van de robuuste regressielijn volgens de methode van Theil is:

$$y = -0{,}0025 + 0{,}0295 \cdot x$$

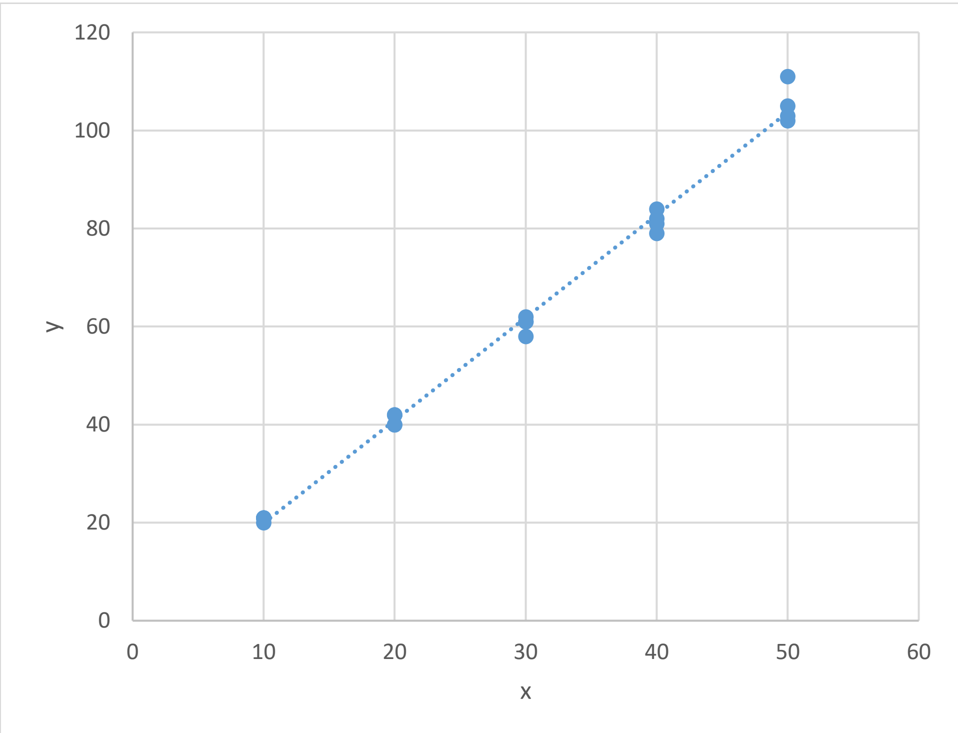
De robuuste regressielijn volgens de methode van Theil is hierna afgebeeld. De verdachte uitbijter ligt duidelijk buiten deze regressielijn.




## Antwoord 5.2

Correcties op antwoorden:
c. weegfactoren $w_i$: 4,4622; 0,3359; 0,1128; 0,0559; 0,0333
d. $b_0 = 0{,}3024$; $b_1 = 2{,}0424$
e. $\hat{x} = 34{,}13 \pm 0{,}012$

Grafiek van de dataset:




Eerst wordt voor elk concentratieniveau *i* de werkelijke standaarddeviatie $s_i$ bepaald van de *y*-waarden van de herhaald gemeten standaarden. Daarna wordt lineaire regressie toegepast op de combinaties van $x_i$ en $s_i$.

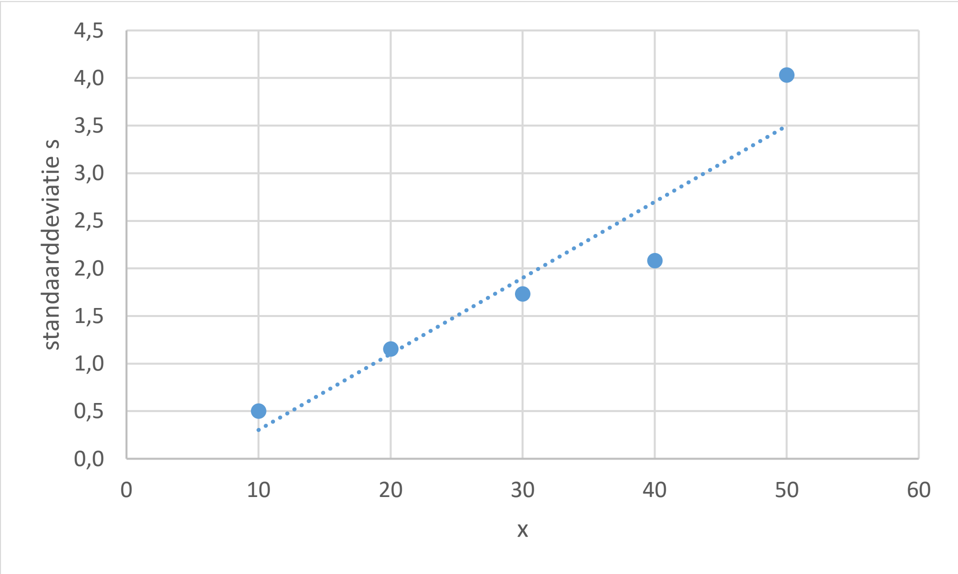
Tip: Gebruik hiervoor de volgende Excel-functies:
$b_0$ =SNIJPUNT(*y*-blok;*x*-blok)
$b_1$ =RICHTING(*y*-blok;*x*-blok)
Het *x*-blok wordt gevormd door de *x*-waarden in de eerste kolom van onderstaande tabel en het *y*-blok door de werkelijke standaarddeviaties $s_i$ in de op een na laatste kolom.

Het resultaat van deze regressie is: $b_0$ = -0,4969; $b_1$ = 0,0799.

Op basis van deze regressieparameters worden de gefitte waarden van de standaarddeviaties $s_i'$ berekend. Zie onderstaande tabel. De waarden voor $s_i'$ worden gebruikt voor de berekening van de weegfactoren bij gewogen lineaire regressie. Zie de tabel voor de berekening van de gewogen regressielijn.

| $x_i$ | y1 | y2 | y3 | y4 | werkelijke sdev $s_i$ | gefitte sdev $s_i'$ |
|---|---|---|---|---|---|---|
| | | | | | | |
| 10 | 20 | 21 | 21 | 21 | 0,500 | 0,3021 |
| 20 | 40 | 40 | 42 | 42 | 1,155 | 1,1010 |
| 30 | 61 | 61 | 62 | 58 | 1,732 | 1,8999 |
| 40 | 84 | 79 | 81 | 82 | 2,082 | 2,6988 |
| 50 | 102 | 103 | 105 | 111 | 4,031 | 3,4978 |

Grafiek van de standaarddeviatie *s* als functie van de concentratie *x*:




De helling $b_{1,w}$ en de as-afsnede $b_{0,w}$ van de gewogen regressielijn kunnen met (5.5) en (5.6) als volgt worden berekend:

$$b_{1,w} = \frac{\sum_{i=1}^{n} w_i x_i y_i - n\bar{x}_w \bar{y}_w}{\sum_{i=1}^{n} w_i x_i^2 - n\bar{x}_w^2} = \frac{7053,555 - 20 \cdot 11,725 \cdot 24,249}{3418,847 - 20 \cdot 11,725^2} = 2,0424$$

$$b_{0,w} = \bar{y}_w - b_{1,w}\bar{x}_w = 24,249 - 2,0424 \cdot 11,725 = 0,3024$$

De vergelijking van de gewogen regressielijn is dus:

$y = 0{,}3024 + 2{,}0424 \cdot x$

waarin *x* de concentratie en *y* de respons is.

Tabel voor de berekening van de gewogen regressielijn:

| $x_i$ | $y_i$ | $s'_i$ | $1/s_i'^2$ | $w_i$ | $w_i x_i$ | $w_i y_i$ | $w_i x_i y_i$ | $w_i x_i^2$ | $\hat{y}_i$ | $w_i(y_i-\hat{y}_i)^2$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 10 | 20 | 0,3021 | 10,960 | 4,4622 | 44,622 | 89,243 | 892,431 | 446,215 | 20,726 | 2,354 |
| 10 | 21 | 0,3021 | 10,960 | 4,4622 | 44,622 | 93,705 | 937,053 | 446,215 | 20,726 | 0,334 |
| 10 | 21 | 0,3021 | 10,960 | 4,4622 | 44,622 | 93,705 | 937,053 | 446,215 | 20,726 | 0,334 |
| 10 | 21 | 0,3021 | 10,960 | 4,4622 | 44,622 | 93,705 | 937,053 | 446,215 | 20,726 | 0,334 |
| 20 | 40 | 1,1010 | 0,825 | 0,3359 | 6,718 | 13,435 | 268,701 | 134,351 | 41,150 | 0,444 |
| 20 | 40 | 1,1010 | 0,825 | 0,3359 | 6,718 | 13,435 | 268,701 | 134,351 | 41,150 | 0,444 |
| 20 | 42 | 1,1010 | 0,825 | 0,3359 | 6,718 | 14,107 | 282,137 | 134,351 | 41,150 | 0,242 |
| 20 | 42 | 1,1010 | 0,825 | 0,3359 | 6,718 | 14,107 | 282,137 | 134,351 | 41,150 | 0,242 |
| 30 | 61 | 1,8999 | 0,277 | 0,1128 | 3,384 | 6,880 | 206,410 | 101,513 | 61,574 | 0,037 |
| 30 | 61 | 1,8999 | 0,277 | 0,1128 | 3,384 | 6,880 | 206,410 | 101,513 | 61,574 | 0,037 |
| 30 | 62 | 1,8999 | 0,277 | 0,1128 | 3,384 | 6,993 | 209,793 | 101,513 | 61,574 | 0,020 |
| 30 | 58 | 1,8999 | 0,277 | 0,1128 | 3,384 | 6,542 | 196,258 | 101,513 | 61,574 | 1,441 |
| 40 | 84 | 2,6988 | 0,137 | 0,0559 | 2,236 | 4,695 | 187,816 | 89,436 | 81,998 | 0,224 |
| 40 | 79 | 2,6988 | 0,137 | 0,0559 | 2,236 | 4,416 | 176,636 | 89,436 | 81,998 | 0,503 |
| 40 | 81 | 2,6988 | 0,137 | 0,0559 | 2,236 | 4,528 | 181,108 | 89,436 | 81,998 | 0,056 |
| 40 | 82 | 2,6988 | 0,137 | 0,0559 | 2,236 | 4,584 | 183,344 | 89,436 | 81,998 | 0,000 |
| 50 | 102 | 3,4978 | 0,082 | 0,0333 | 1,664 | 3,394 | 169,721 | 83,197 | 102,422 | 0,006 |
| 50 | 103 | 3,4978 | 0,082 | 0,0333 | 1,664 | 3,428 | 171,385 | 83,197 | 102,422 | 0,011 |
| 50 | 105 | 3,4978 | 0,082 | 0,0333 | 1,664 | 3,494 | 174,713 | 83,197 | 102,422 | 0,221 |
| 50 | 111 | 3,4978 | 0,082 | 0,0333 | 1,664 | 3,694 | 184,697 | 83,197 | 102,422 | 2,449 |
| | | | | | | | | | | |
| som | | | 49,123 | 20,0000 | 234,491 | 484,971 | 7053,555 | 3418,847 | | 9,735 |
| gemiddelde | | | 2,456 | 1,0000 | 11,725 | 24,249 | 352,678 | 170,942 | | 0,487 |

De standaarddeviatie $s_{r,w}$ van de gewogen regressielijn berekend met (5.9) is:

$$s_{r,w} = \sqrt{\frac{\sum_{i=1}^{n} w_i(y_i - \hat{y}_i)^2}{n-2}} = \sqrt{\frac{9{,}735}{20-2}} = 0{,}735$$

De geschatte concentratie $\hat{x}_0$ bij een meetsignaal $y_0$ = 70 in één monster is:

$$\hat{x}_0 = \frac{y_0 - b_{0,w}}{b_{1,w}} = \frac{70 - 0{,}3024}{2{,}0424} = 34{,}13$$

Het bijbehorende betrouwbaarheidsinterval wordt berekend met (5.10), $m$ = 1 en de tweezijdige $t_{(0,05;18)}$ = 2,101.

De weegfactor $w_0$ die hoort bij $\hat{x}_0$ wordt als volgt berekend met behulp van de regressielijn voor de standaarddeviatie met $b_0$ = -0,4969 en $b_1$ = 0,0799. Voor $\hat{x}_0 = 34{,}13$ is $s'_0 = -0{,}4969 + 0{,}0799 \cdot 34{,}13 = 2{,}229$; $1/s_0'^2 = 1/2{,}229^2 = 0{,}2012$ en met (5.3):

$$w_0 = \frac{1/s_0^2}{\left(\sum_{i=1}^{n} 1/s_{y_i}^2\right)/n} = \frac{0{,}2012}{49{,}123/20} = 0{,}082$$

Berekening $\mathrm{BI}(\hat{x}_0)$ met (5.10):

$$\mathrm{BI}(\hat{x}_0) = \hat{x}_0 \pm \frac{t_{(0,05;n-p)} s_{r,w}}{b_{1,w}} \cdot \sqrt{\frac{1}{w_0 m} + \frac{1}{n} + \frac{(y_0 - \bar{y}_w)^2}{b_{1,w}^2 (\sum w_i x_i^2 - n\bar{x}_w^2)}}$$

$$\mathrm{BI}(\hat{x}_0) = 34{,}13 \pm \frac{2{,}101 \cdot 0{,}735}{2{,}0424} \cdot \sqrt{\frac{1}{0{,}082 \cdot 1} + \frac{1}{20} + \frac{(70 - 24{,}249)^2}{2{,}0424^2 (3418{,}847 - 20 \cdot 11{,}725^2)}}$$

$$\mathrm{BI}(\hat{x}_0) = 34{,}13 \pm 0{,}012$$

## Antwoord 5.3

| i | $x_{i1}$ | $x_{i2}$ | $y_{i1}$ | $y_{i2}$ | | $\bar{x}_i$ | $\bar{y}_i$ |
|---|---|---|---|---|---|---|---|
| 1 | 13,80 | 13,80 | 12,50 | 12,10 | | 13,80 | 12,30 |
| 2 | 15,40 | 15,40 | 18,20 | 14,60 | | 15,40 | 16,40 |
| 3 | 16,00 | 14,90 | 15,90 | 16,50 | | 15,45 | 16,20 |
| 4 | 22,40 | 24,10 | 22,70 | 22,40 | | 23,25 | 22,55 |
| 5 | 22,90 | 22,50 | 20,00 | 22,80 | | 22,70 | 21,40 |
| 6 | 26,10 | 26,40 | 24,80 | 25,80 | | 26,25 | 25,30 |
| 7 | 29,90 | 28,60 | 28,50 | 30,30 | | 29,25 | 29,40 |
| 8 | 31,30 | 30,10 | 29,10 | 30,60 | | 30,70 | 29,85 |
| 9 | 33,00 | 33,80 | 33,10 | 31,90 | | 33,40 | 32,50 |
| 10 | 38,90 | 40,70 | 40,10 | 39,70 | | 39,80 | 39,90 |

Voor elke *i* wordt het gemiddelde van $x_1$ en $y_1$ uit de voorgaande tabel ingevuld als $x_\mathrm{i}$ en $y_\mathrm{i}$ in de volgende tabel.

| i | $x_i$ | $y_i$ | $(x_i - \bar{x})$ | $(x_i - \bar{x})^2$ | $(y_i - \bar{y})$ | $(y_i - \bar{y})^2$ | $(x_i - \bar{x})(y_i - \bar{y})$ |
|---|---|---|---|---|---|---|---|
| 1 | 13,80 | 12,30 | -11,200 | 125,4400 | -12,2800 | 150,7984 | 137,5360 |
| 2 | 15,40 | 16,40 | -9,600 | 92,1600 | -8,1800 | 66,9124 | 78,5280 |
| 3 | 15,45 | 16,20 | -9,550 | 91,2025 | -8,3800 | 70,2244 | 80,0290 |
| 4 | 23,25 | 22,55 | -1,750 | 3,0625 | -2,0300 | 4,1209 | 3,5525 |
| 5 | 22,70 | 21,40 | -2,300 | 5,2900 | -3,1800 | 10,1124 | 7,3140 |
| 6 | 26,25 | 25,30 | 1,250 | 1,5625 | 0,7200 | 0,5184 | 0,9000 |
| 7 | 29,25 | 29,40 | 4,250 | 18,0625 | 4,8200 | 23,2324 | 20,4850 |
| 8 | 30,70 | 29,85 | 5,700 | 32,4900 | 5,2700 | 27,7729 | 30,0390 |
| 9 | 33,40 | 32,50 | 8,400 | 70,5600 | 7,9200 | 62,7264 | 66,5280 |
| 10 | 39,80 | 39,90 | 14,800 | 219,0400 | 15,3200 | 234,7024 | 226,7360 |
| | | | | | | | |
| som | 250,00 | 245,80 | | 658,8700 | | 651,1210 | 651,6475 |
| gem | 25,00 | 24,58 | | $SS_{xx}$ | | $SS_{yy}$ | $SS_{xy}$ |

De helling van de orthogonale regressielijn kan worden berekend met (5.11):

$$b_1 = \frac{SS_{\mathrm{yy}} - SS_{\mathrm{xx}} + \sqrt{\left(SS_{\mathrm{yy}} - SS_{\mathrm{xx}}\right)^2 + 4SS_{\mathrm{xy}}^2}}{2SS_{\mathrm{xy}}}$$

$$b_1 = \frac{651{,}1210 - 658{,}8700 + \sqrt{(651{,}1210 - 658{,}8700)^2 + 4 \cdot 651{,}6475^2}}{2 \cdot 651{,}6475} = 0{,}9941$$

De as-afsnede van de orthogonale regressielijn kan worden berekend met (5.12):
$b_0 = \bar{y} - b_1\bar{x}$

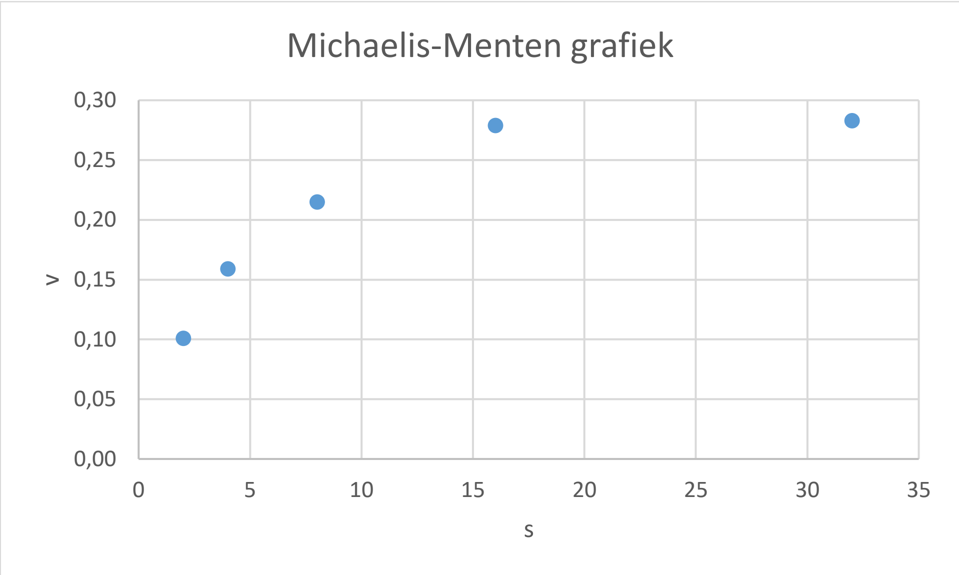
$$b_0 = 24{,}58 - 0{,}9941 \cdot 25{,}00 = -0{,}2718$$

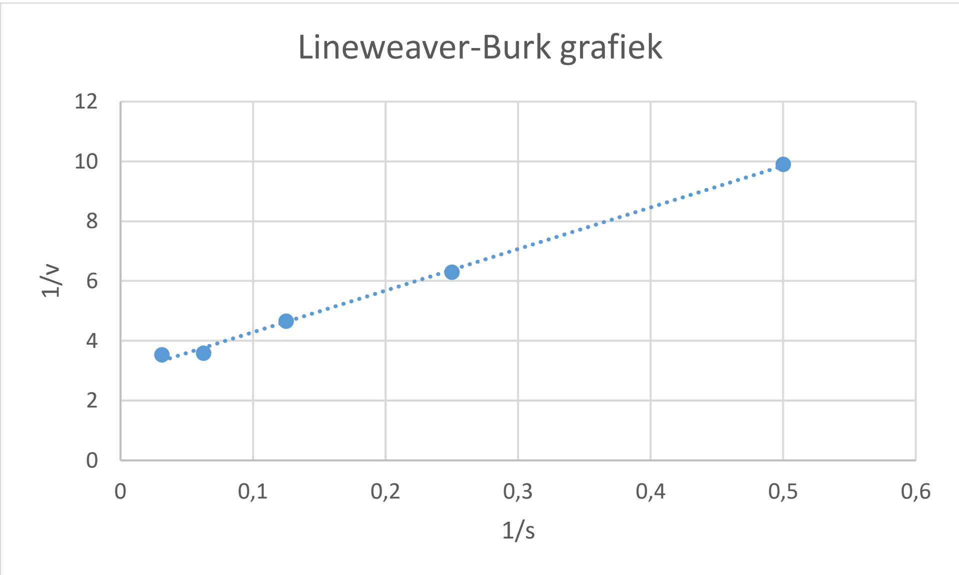
De vergelijking van de orthogonale regressielijn is dus:

$$y = -0{,}272 + 0{,}994 \cdot x$$

## Antwoord 5.4

| no | [S] | v | 1/[S] | 1/v |
|---|---|---|---|---|
| 1 | 2 | 0,101 | 0,5 | 9,9010 |
| 2 | 4 | 0,159 | 0,25 | 6,2893 |
| 3 | 8 | 0,215 | 0,125 | 4,6512 |
| 4 | 16 | 0,279 | 0,0625 | 3,5842 |
| 5 | 32 | 0,283 | 0,03125 | 3,5336 |



Regressie voor de Lineweaver-Burk grafiek

Tip: Gebruik hiervoor de volgende Excel-functies:
$b_0$ =SNIJPUNT(*y*-blok;*x*-blok)
$b_1$ =RICHTING(*y*-blok;*x*-blok)
Het *x*-blok wordt gevormd door de 1/[*S*]-waarden van voorgaande tabel en het *y*-blok door de 1/*v* waarden.

Het resultaat is: $b_0$ = 2,8965; $b_1$ = 13,9117.

$$v_{\max} = \frac{1}{b_0} = \frac{1}{2{,}8965} = 0{,}345$$

$$K_{\mathrm{M}} = \frac{b_1}{b_0} = \frac{13{,}9117}{2{,}8965} = 4{,}803$$

Niet-lineaire regressie met de Oplosser in Excel

| enzym [S] | snelheid v | snelheid geschat | kwadraat residu | | |
|---|---|---|---|---|---|
| 2,0 | 0,10 | 0,105 | 1,82E-05 | v_max= | 0,334 |
| 4,0 | 0,16 | 0,160 | 1,23E-06 | KM= | 4,350 |
| 8,0 | 0,22 | 0,217 | 2,27E-06 | | |
| 16,0 | 0,28 | 0,263 | 2,63E-04 | | |
| 32,0 | 0,28 | 0,294 | 1,26E-04 | | |
| | | kwadratensom | 4,11E-04 | | |

## Antwoord 5.5

Correctie op tekst in de opgave
‘Opgave 1’ en ‘Opgave 2’ moeten zijn ‘opgave 4.2’.

Correctie op de antwoorden bij opgave 5.5.

a. Zie opgave 4.2
b. Zie opgave 4.3

**Antwoord 5.5.a**: Matrixregressie met het eerstegraads model voor de kalibratieset van opgave 4.2:

De $\boldsymbol{y}$-vector en $\boldsymbol{X}$-matrix in het model $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{b}$ zijn:

$$\boldsymbol{y} = \begin{bmatrix} 0{,}075 \\ 0{,}182 \\ 0{,}248 \\ 0{,}357 \\ 0{,}401 \end{bmatrix} \qquad \boldsymbol{X} = \begin{bmatrix} 1 & 0{,}005 \\ 1 & 0{,}010 \\ 1 & 0{,}015 \\ 1 & 0{,}020 \\ 1 & 0{,}025 \end{bmatrix}$$

De modelparameters kunnen als volgt worden berekend met (5.19):

$$\boldsymbol{X^T X} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0{,}005 & 0{,}010 & 0{,}015 & 0{,}020 & 0{,}025 \end{bmatrix} \begin{bmatrix} 1 & 0{,}005 \\ 1 & 0{,}010 \\ 1 & 0{,}015 \\ 1 & 0{,}020 \\ 1 & 0{,}025 \end{bmatrix} = \begin{bmatrix} 5 & 0{,}075 \\ 0{,}075 & 0{,}001375 \end{bmatrix}$$

$$(\boldsymbol{X^T X})^{-1} = \begin{bmatrix} 1{,}1000 & -60 \\ -60 & 4000 \end{bmatrix}$$

$$\boldsymbol{\hat{b}} = (\boldsymbol{X^T X})^{-1} \boldsymbol{X^T y}$$

$$\boldsymbol{\hat{b}} = \begin{bmatrix} 1{,}1000 & -60 \\ -60 & 4000 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0{,}005 & 0{,}010 & 0{,}015 & 0{,}020 & 0{,}025 \end{bmatrix} \cdot \begin{bmatrix} 0{,}075 \\ 0{,}182 \\ 0{,}248 \\ 0{,}357 \\ 0{,}401 \end{bmatrix} = \begin{bmatrix} 0{,}0045 \\ 16{,}5400 \end{bmatrix}$$

$b_0$ = 0,0045 en $b_1$ = 16,54.
De vector met geschatte meetwaarden $\boldsymbol{\hat{y}}$ wordt berekend met (5.20):

$$\boldsymbol{\hat{y}} = \boldsymbol{X\hat{b}} = \begin{bmatrix} 1 & 0{,}005 \\ 1 & 0{,}010 \\ 1 & 0{,}015 \\ 1 & 0{,}020 \\ 1 & 0{,}025 \end{bmatrix} \cdot \begin{bmatrix} 0{,}0045 \\ 16{,}5400 \end{bmatrix} = \begin{bmatrix} 0{,}0872 \\ 0{,}1699 \\ 0{,}2526 \\ 0{,}3353 \\ 0{,}4180 \end{bmatrix}$$

De residuen worden berekend met (5.21):

$$\boldsymbol{r} = \boldsymbol{y} - \boldsymbol{\hat{y}} = \begin{bmatrix} 0{,}075 \\ 0{,}182 \\ 0{,}248 \\ 0{,}357 \\ 0{,}401 \end{bmatrix} - \begin{bmatrix} 0{,}0872 \\ 0{,}1699 \\ 0{,}2526 \\ 0{,}3353 \\ 0{,}4180 \end{bmatrix} = \begin{bmatrix} -0{,}0122 \\ 0{,}0121 \\ -0{,}0046 \\ 0{,}0217 \\ -0{,}0170 \end{bmatrix}$$

De variantie van de residuen wordt berekend met (5.23):

$$s_r^2 = \frac{1}{(5-2)} \begin{bmatrix} -0{,}0122 & 0{,}0121 & -0{,}0046 & 0{,}0217 & -0{,}0170 \end{bmatrix} \cdot \begin{bmatrix} -0{,}0122 \\ 0{,}0121 \\ -0{,}0046 \\ 0{,}0217 \\ -0{,}0170 \end{bmatrix} = 3{,}59 \cdot 10^{-4}$$

De variantie-covariantiematrix wordt berekend met (5.26):

$$\boldsymbol{V}(\boldsymbol{b}) = s_r^2 (\boldsymbol{X^T X})^{-1} = 3{,}59 \cdot 10^{-4} \cdot \begin{bmatrix} 1{,}1000 & -60 \\ -60 & 4000 \end{bmatrix} = \begin{bmatrix} 3{,}95 \cdot 10^{-4} & -2{,}15 \cdot 10^{-2} \\ -2{,}15 \cdot 10^{-2} & 1{,}435 \end{bmatrix}$$

De standaarddeviaties van de geschatte modelparameters kunnen worden berekend uit de diagonaalelementen van de variantie-covariantiematrix:

$$s_{b_0} = \sqrt{3{,}95 \cdot 10^{-4}} = 1{,}99 \cdot 10^{-2}$$
$$s_{b_1} = \sqrt{1{,}435} = 1{,}20$$

De betrouwbaarheidsintervallen van de modelparameters kunnen worden berekend door deze waarden in te vullen in (5.27) met $t_{(0,05;3)}$ = 3.182, zie tabel 3 in bijlage 1.

$\mathrm{BI}(b_0) = b_0 \pm t_{(0{,}05;n-p)} s_{b_0} = 0{,}0045 \pm 3{,}182 \cdot 1{,}99 \cdot 10^{-2} = 0{,}0045 \pm 6{,}32 \cdot 10^{-2}$

$\mathrm{BI}(b_1) = b_1 \pm t_{(0{,}05;n-p)} s_{b_1} = 16{,}54 \pm 3{,}182 \cdot 1{,}20 = 16{,}54 \pm 3{,}81$

**Antwoord 5.5.b**: Matrixregressie met het tweedegraads model voor de kalibratieset van opgave 4.2:

De $\boldsymbol{y}$-vector en $\boldsymbol{X}$-matrix in het model $\boldsymbol{y} = \boldsymbol{Xb}$ zijn:

$$\boldsymbol{y} = \begin{bmatrix} 0{,}075 \\ 0{,}182 \\ 0{,}248 \\ 0{,}357 \\ 0{,}401 \end{bmatrix} \qquad \boldsymbol{X} = \begin{bmatrix} 1 & 5{,}0 \cdot 10^{-3} & 2{,}50 \cdot 10^{-5} \\ 1 & 1{,}0 \cdot 10^{-2} & 1{,}00 \cdot 10^{-4} \\ 1 & 1{,}5 \cdot 10^{-2} & 2{,}25 \cdot 10^{-4} \\ 1 & 2{,}0 \cdot 10^{-2} & 4{,}00 \cdot 10^{-4} \\ 1 & 2{,}5 \cdot 10^{-2} & 6{,}25 \cdot 10^{-4} \end{bmatrix}$$

De modelparameters kunnen worden berekend met (5.19):

$$\hat{\boldsymbol{b}} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y}$$

waarin,

$$(\boldsymbol{X}^T\boldsymbol{X})^{-1} = \begin{bmatrix} 4{,}60 & -6{,}600 \cdot 10^2 & 2{,}000 \cdot 10^4 \\ -6{,}600 \cdot 10^2 & 1{,}069 \cdot 10^5 & -3{,}429 \cdot 10^6 \\ 2{,}000 \cdot 10^4 & -3{,}429 \cdot 10^6 & 1{,}143 \cdot 10^8 \end{bmatrix}$$

$$\hat{\boldsymbol{b}} = \begin{bmatrix} -0{,}0370 \\ 23{,}6543 \\ -237{,}1429 \end{bmatrix}$$

$b_0$ = -0,0370, $b_1$ = 23,6543 en $b_2$ = -237,1429.

De vector met geschatte meetwaarden $\hat{\boldsymbol{y}}$ wordt berekend met (5.20):

$$\hat{\boldsymbol{y}} = \boldsymbol{X}\hat{\boldsymbol{b}} = \begin{bmatrix} 1 & 5{,}0 \cdot 10^{-3} & 2{,}50 \cdot 10^{-5} \\ 1 & 1{,}0 \cdot 10^{-2} & 1{,}00 \cdot 10^{-4} \\ 1 & 1{,}5 \cdot 10^{-2} & 2{,}25 \cdot 10^{-4} \\ 1 & 2{,}0 \cdot 10^{-2} & 4{,}00 \cdot 10^{-4} \\ 1 & 2{,}5 \cdot 10^{-2} & 6{,}25 \cdot 10^{-4} \end{bmatrix} \cdot \begin{bmatrix} -0{,}0370 \\ 23{,}6543 \\ -237{,}1429 \end{bmatrix} = \begin{bmatrix} 0{,}0753 \\ 0{,}1758 \\ 0{,}2645 \\ 0{,}3412 \\ 0{,}4061 \end{bmatrix}$$

De residuen worden berekend met (5.21):

$$\boldsymbol{r} = \boldsymbol{y} - \hat{\boldsymbol{y}} = \begin{bmatrix} 0{,}075 \\ 0{,}182 \\ 0{,}248 \\ 0{,}357 \\ 0{,}401 \end{bmatrix} - \begin{bmatrix} 0{,}0753 \\ 0{,}1758 \\ 0{,}2645 \\ 0{,}3412 \\ 0{,}4061 \end{bmatrix} = \begin{bmatrix} -0{,}0003 \\ 0{,}0062 \\ -0{,}0165 \\ 0{,}0158 \\ -0{,}0051 \end{bmatrix}$$

De variantie van de residuen wordt berekend met (5.23):

$$s_r^2 = \frac{1}{(5-3)}\begin{bmatrix}-0{,}0003 & 0{,}0062 & -0{,}0165 & 0{,}0158 & -0{,}0051\end{bmatrix} \cdot \begin{bmatrix}-0{,}0003\\ 0{,}0062\\ -0{,}0165\\ 0{,}0158\\ -0{,}0051\end{bmatrix} = 2{,}92 \cdot 10^{-4}$$

De variantie-covariantiematrix wordt berekend met (5.26):

$$\boldsymbol{V}(\boldsymbol{b}) = s_r^2(\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{X})^{-1} = 2{,}92 \cdot 10^{-4} \cdot \begin{bmatrix} 4{,}60 & -6{,}600 \cdot 10^2 & 2{,}000 \cdot 10^4 \\ -6{,}600 \cdot 10^2 & 1{,}069 \cdot 10^5 & -3{,}429 \cdot 10^6 \\ 2{,}000 \cdot 10^4 & -3{,}429 \cdot 10^6 & 1{,}143 \cdot 10^8 \end{bmatrix}$$
$$= \begin{bmatrix} 1{,}344 \cdot 10^{-3} & -1{,}928 \cdot 10^{-1} & 5{,}842 \\ -1{,}928 \cdot 10^{-1} & 3{,}121 \cdot 10^1 & -1{,}002 \cdot 10^3 \\ 5{,}842 & -1{,}002 \cdot 10^3 & 3{,}338 \cdot 10^4 \end{bmatrix}$$

De standaarddeviaties van de geschatte modelparameters kunnen worden berekend uit de diagonaalelementen van de variantie-covariantiematrix:

$$s_{b_0} = \sqrt{1{,}344 \cdot 10^{-3}} = 0{,}0367$$
$$s_{b_1} = \sqrt{3{,}121 \cdot 10^1} = 5{,}587$$
$$s_{b_2} = \sqrt{3{,}338 \cdot 10^4} = 182{,}71$$

De betrouwbaarheidsintervallen van de modelparameters kunnen worden berekend door deze waarden in te vullen in (5.27) met $t_{(0,05;2)} = 4{,}303$, zie tabel 3 in bijlage 1.

$$\mathrm{BI}(b_0) = b_0 \pm t_{(0,05;n-p)}s_{b_0} = -0{,}0370 \pm 4{,}303 \cdot 0{,}0367 = -0{,}0370 \pm 0{,}1577$$

$$\mathrm{BI}(b_1) = b_1 \pm t_{(0,05;n-p)}s_{b_1} = 23{,}65 \pm 4{,}303 \cdot 5{,}587 = 23{,}65 \pm 24{,}03$$

$$\mathrm{BI}(b_2) = b_2 \pm t_{(0,05;n-p)}s_{b_2} = -237 \pm 4{,}303 \cdot 182{,}71 = -237 \pm 786$$