Lineær regression = bedste rette linje gennem et antal punkter

LinReg fra Gym-pakken

Med *Gym*-pakken kan vi få beregnet og tegnet den bedste rette linje:

```
> restart
> with(Gym) :
```

Dataene indskrives i 2 lister:

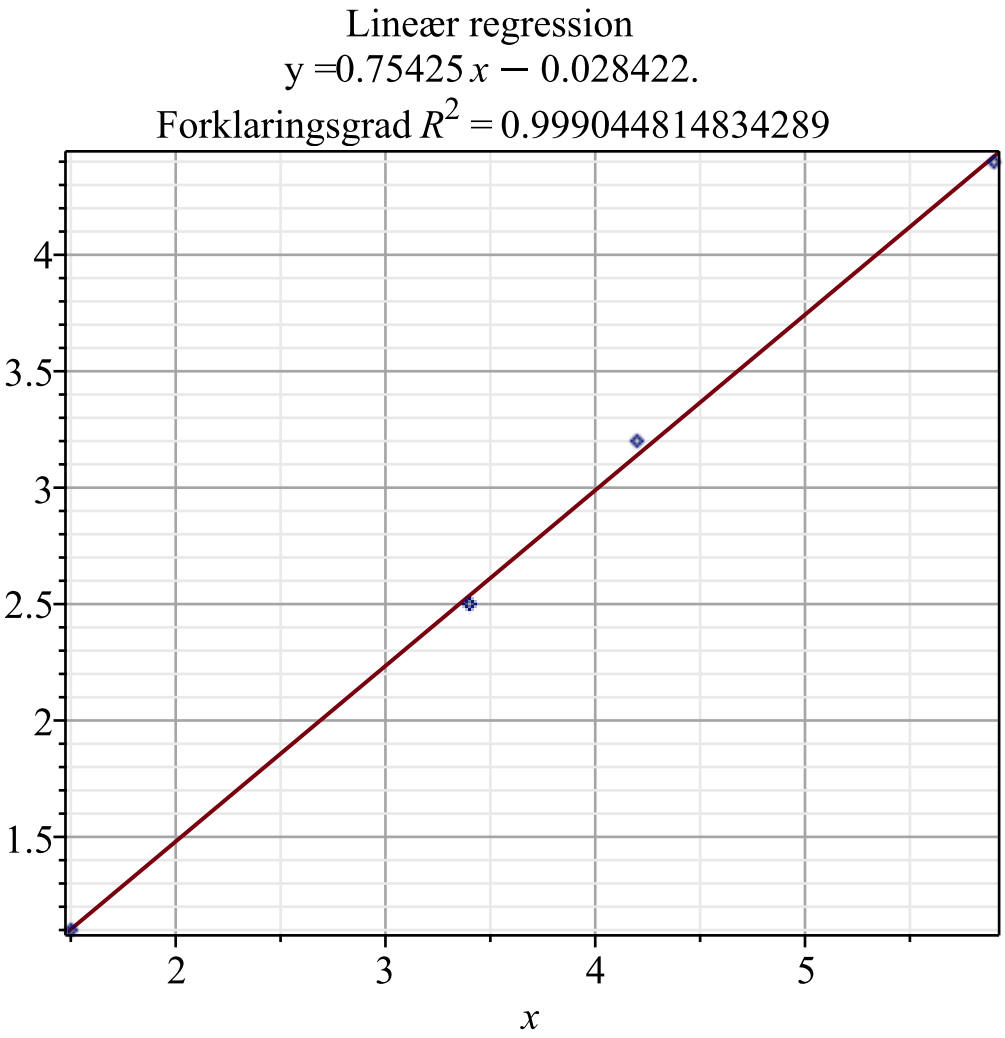
```
> X := [1.5, 3.4, 4.2, 5.9]
```

$$X := [1.5, 3.4, 4.2, 5.9] \quad \textbf{(1.1)}$$

```
> Y := [1.1, 2.5, 3.2, 4.4]
```

$$Y := [1.1, 2.5, 3.2, 4.4] \quad \textbf{(1.2)}$$

```
> LinReg(X, Y)
```



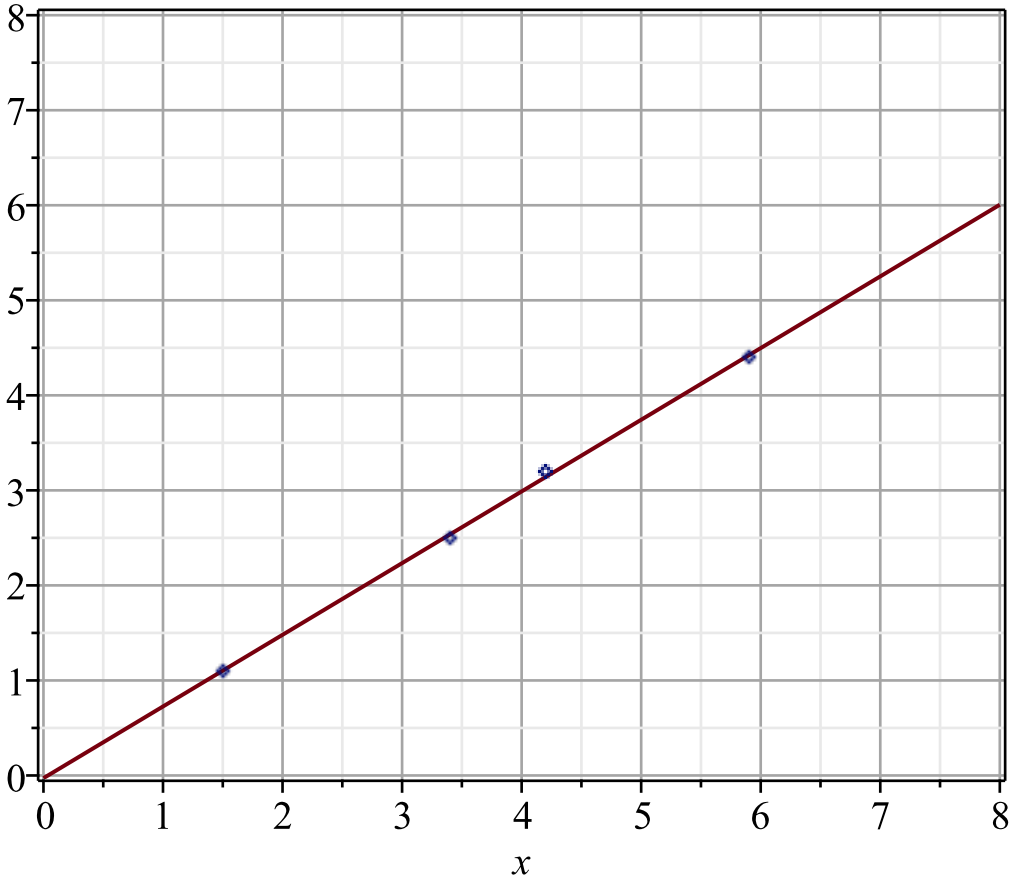
Hvis man ønsker en graf, hvor man kan se origo, dvs. (0,0), så kan man gøre sådan:

```
> LinReg(X, Y, view = [0..8, 0..8])
```

Lineær regression

$y = 0.75425\,x - 0.028422.$

Forklaringsgrad $R^2 = 0.999044814834289$



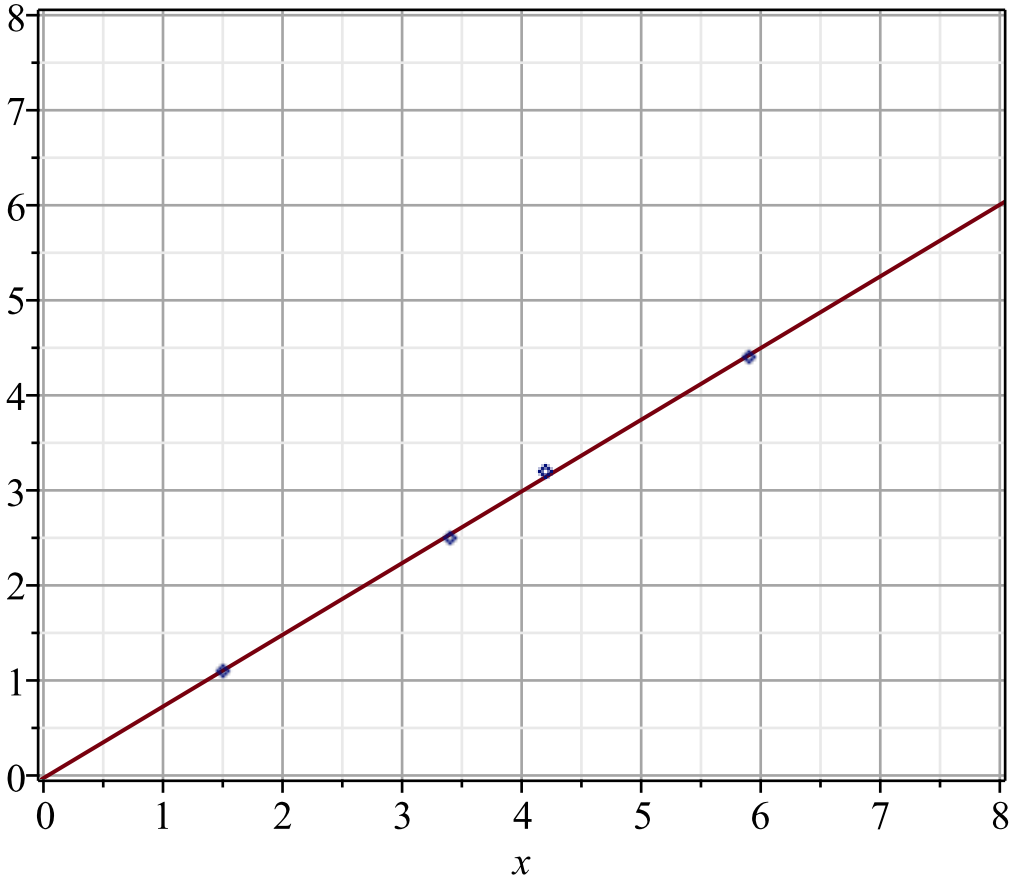
Eller, hvis man selv vil styre det:

```
> with(plots):
  Graf := LinReg(X, Y):
  display(Graf, view = [0..8, 0..8])
```

Lineær regression

$y = 0.75425\,x - 0.028422.$

Forklaringsgrad $R^2 = 0.999044814834289$



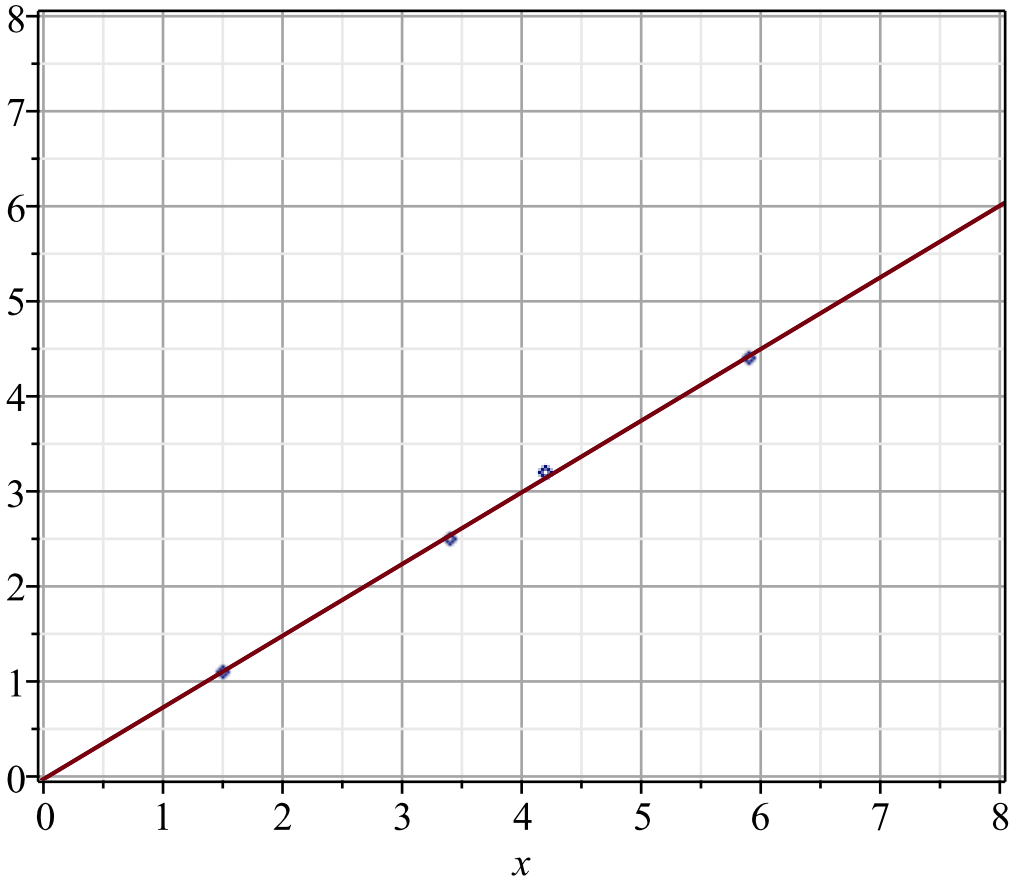
Her styrer man det hele selv:

```
> with(plots) :
  Graf := LinReg(X, Y) :
  Funktion := unapply(LinReg(X, Y, x), x) :
  Linje := plot(Funktion) :
  display(Graf, Linje, view = [0..8, 0..8])
```

Lineær regression

$y = 0.75425\,x - 0.028422.$

Forklaringsgrad $R^2 = 0.999044814834289$



Regressionsfunktionen og beregning af x eller y og hvordan man suger a og b ud

Med *LinReg(X,Y,x)* kan man få funktionsudtrykket:

> $LinReg(X, Y, x)$

$$0.754245754245754\,x - 0.0284215784215784 \qquad \textbf{(2.1)}$$

Det kan man gemme i en funktion $f(x)$.
Hertil anvendes den driftssikre *unapply* :

> $f := unapply(LinReg(X, Y, x), x)$

$$f := x \to 0.7542457542457542\,x - 0.02842157842157844 \qquad \textbf{(2.2)}$$

Hvis $x = 5$ er givet, hvad er så y ?

> $f(5)$

$$3.74280719280719 \qquad \textbf{(2.3)}$$

Hvis $y = 4$ er givet, hvad er så x ?

> $solve(f(x) = 4, x)$

$$5.340993377 \qquad \textbf{(2.4)}$$

Hvordan kan man suge a og b ud af regressionsfunktionen $a \cdot x + b$?
- så man slipper for copy-paste!

> $b := f(0)$

$$b := -0.0284215784215784 \qquad \textbf{(2.5)}$$

> $a := f(1) - f(0)$

$$a := 0.754245754245754 \qquad \textbf{(2.6)}$$