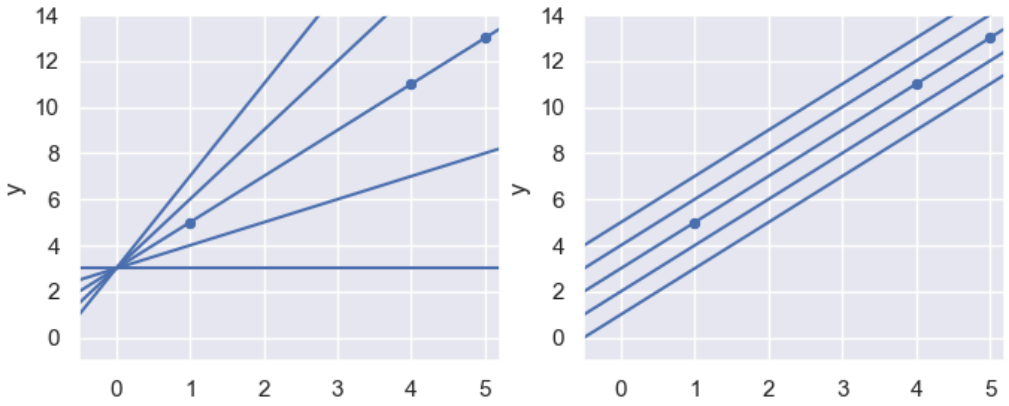
# Exercise 1 - Mean squared error

Consider the three points $(1, 5)$, $(4, 11)$ and $(5, 13)$:




- The left graphics shows the lines $f(x) = 3 + 0 \cdot x$, $f(x) = 3 + 1 \cdot x$, $f(x) = 3 + 2 \cdot x$, $f(x) = 3 + 3 \cdot x$, $f(x) = 3 + 4 \cdot x$. Calculate the mean-squared error for each straight line. Then plot the slope and the resulting mean-squared error in a common coordinate system. Which functional form do you recognize?
- The right graphics shows the lines $f(x) = 1 + 2 \cdot x$, $f(x) = 3 + 2 \cdot x$, $f(x) = 4 + 2 \cdot x$, $f(x) = 5 + 2 \cdot x$, $f(x) = 6 + 2 \cdot x$. Calculate the mean-squared error for each straight line. Then plot the intercept and the resulting mean-squared error in a common coordinate system. Which functional form do you recognize?

# Exercise 2 - USB sticks

The following table shows current market prices for USB flash drives of different sizes.

| **Capacity [GB]** | **Price [€]** |
|---|---|
| 16 | 6 |
| 32 | 6.5 |
| 64 | 8.5 |
| 128 | 14 |
| 256 | 28 |

1. Why is it justified to model the prices using linear regression?
2. Compute the optimal regression line manually, i.e. with a calculator, using the formula from the lecture.
3. What interpretation do the parameters in this example allow?

Have a closer look at the following webpages:

https://scikit-learn.org/stable/modules/linear_model.html (only Section 1.1.1; but take in an impression of what else is possible!)

https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LinearRegression.html