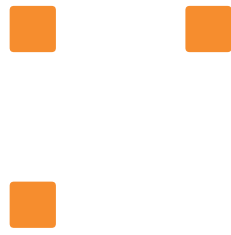
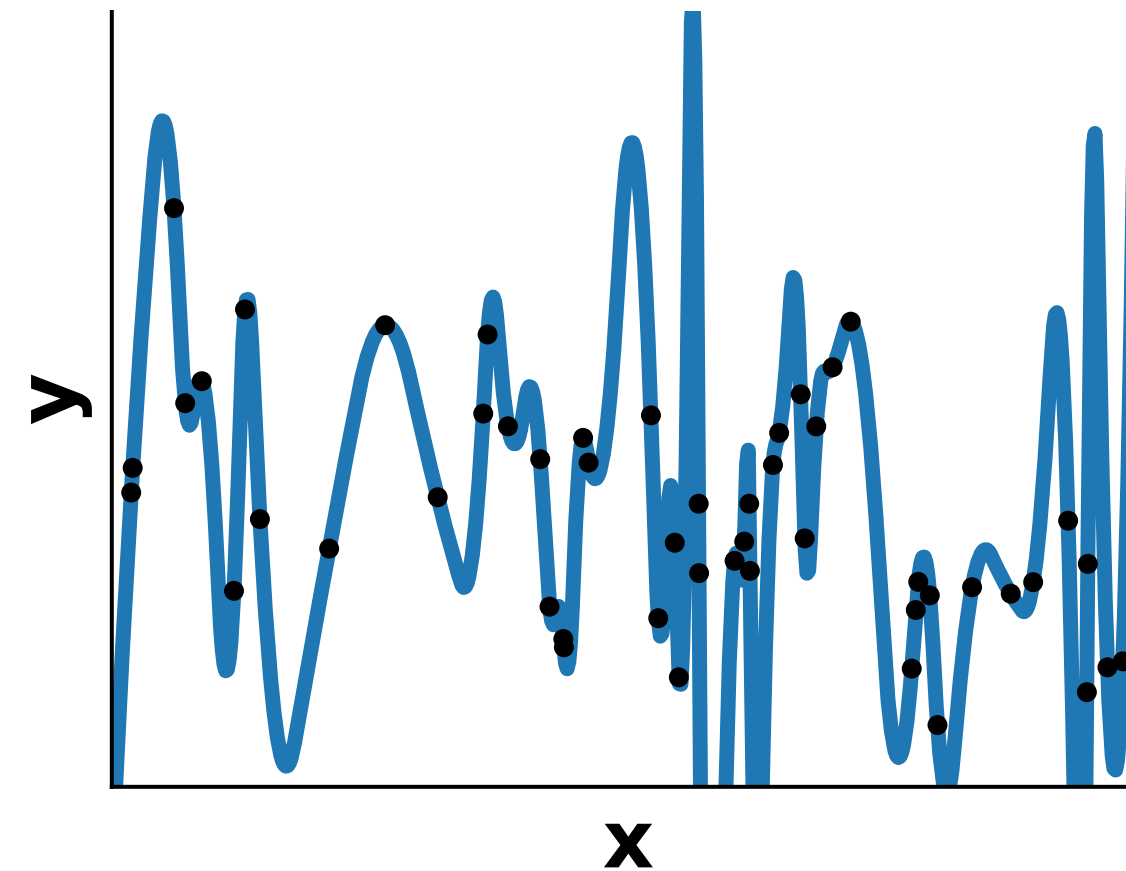
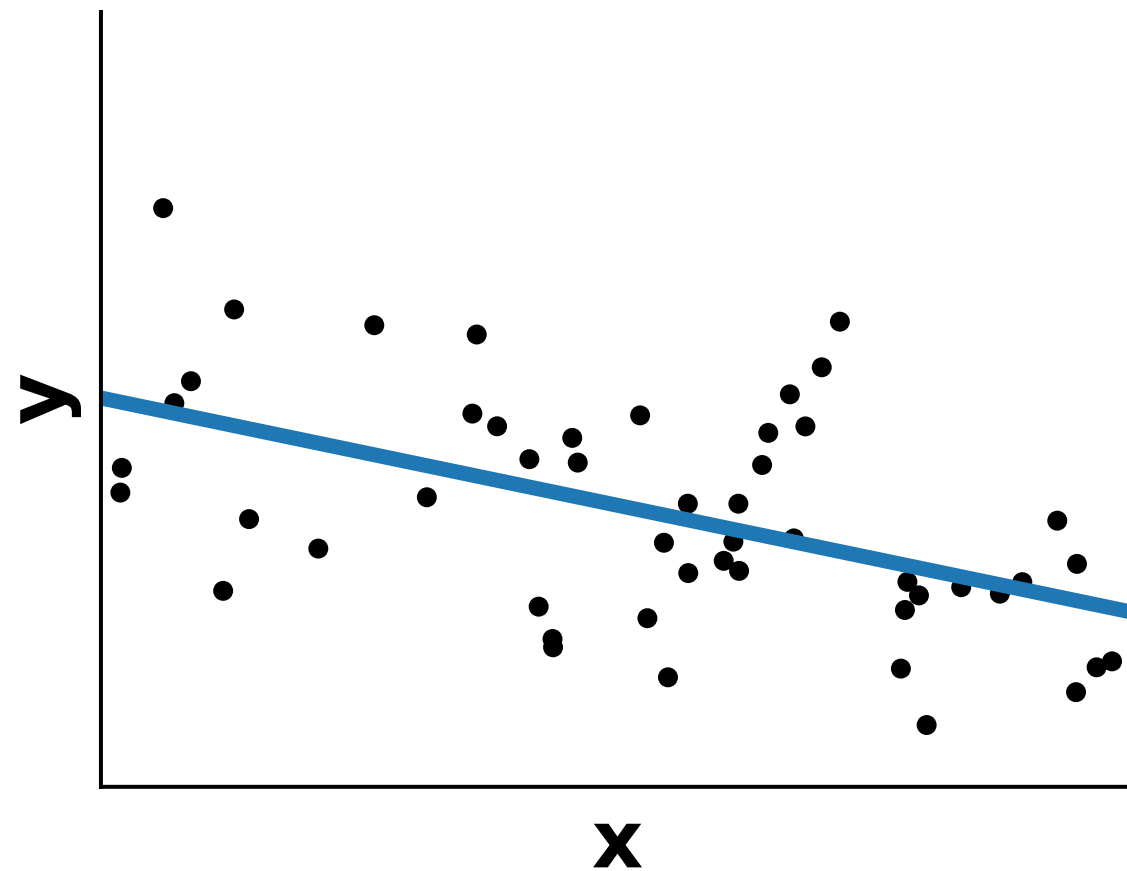


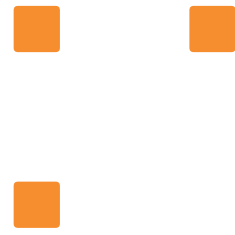
Overfitting and underfitting

Understand when and why a model does or does not generalize well on unseen data.

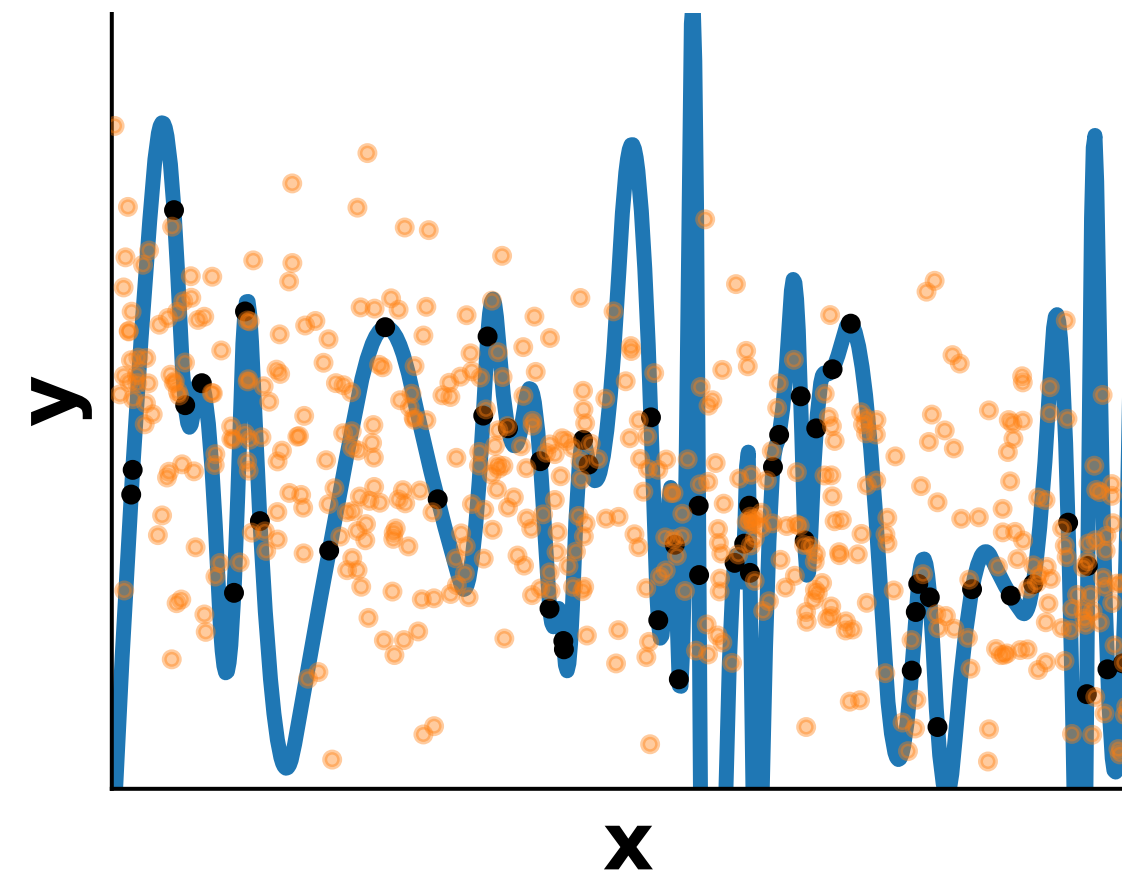
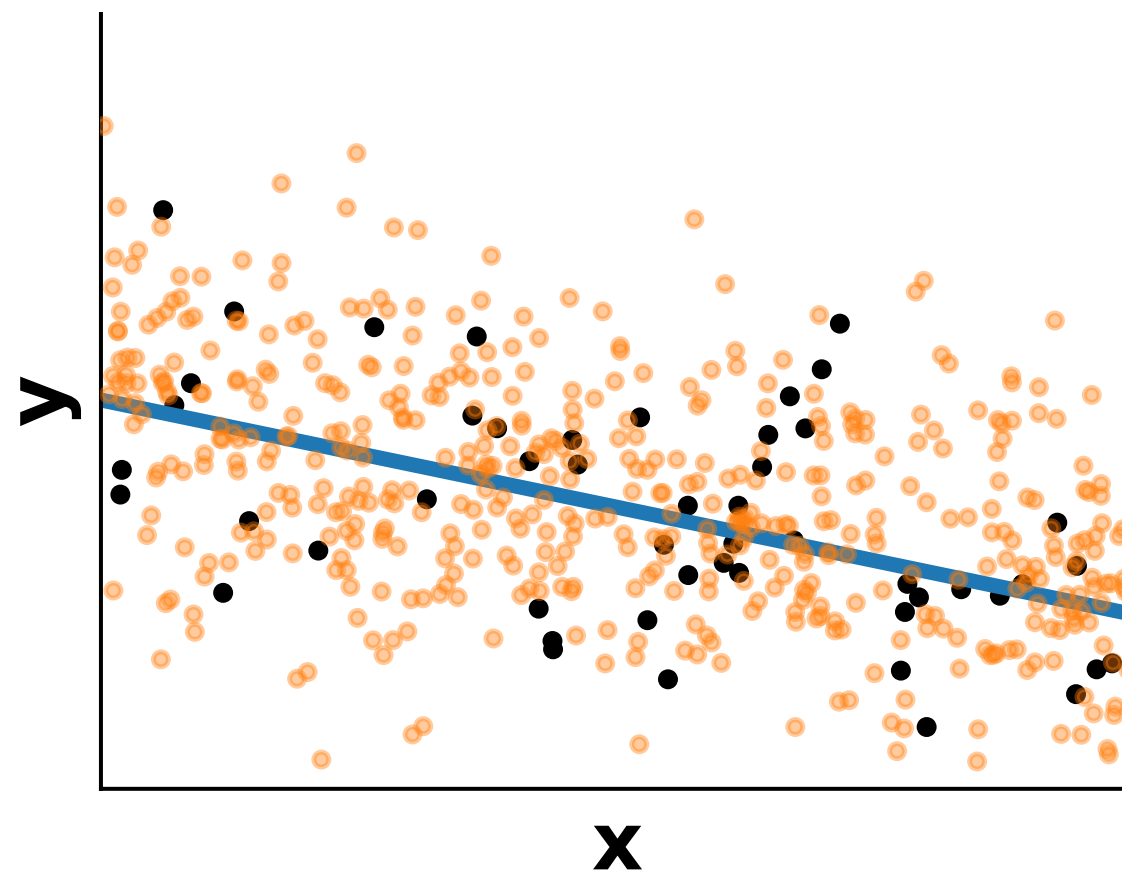


Which data fit do you prefer?

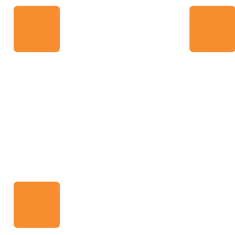




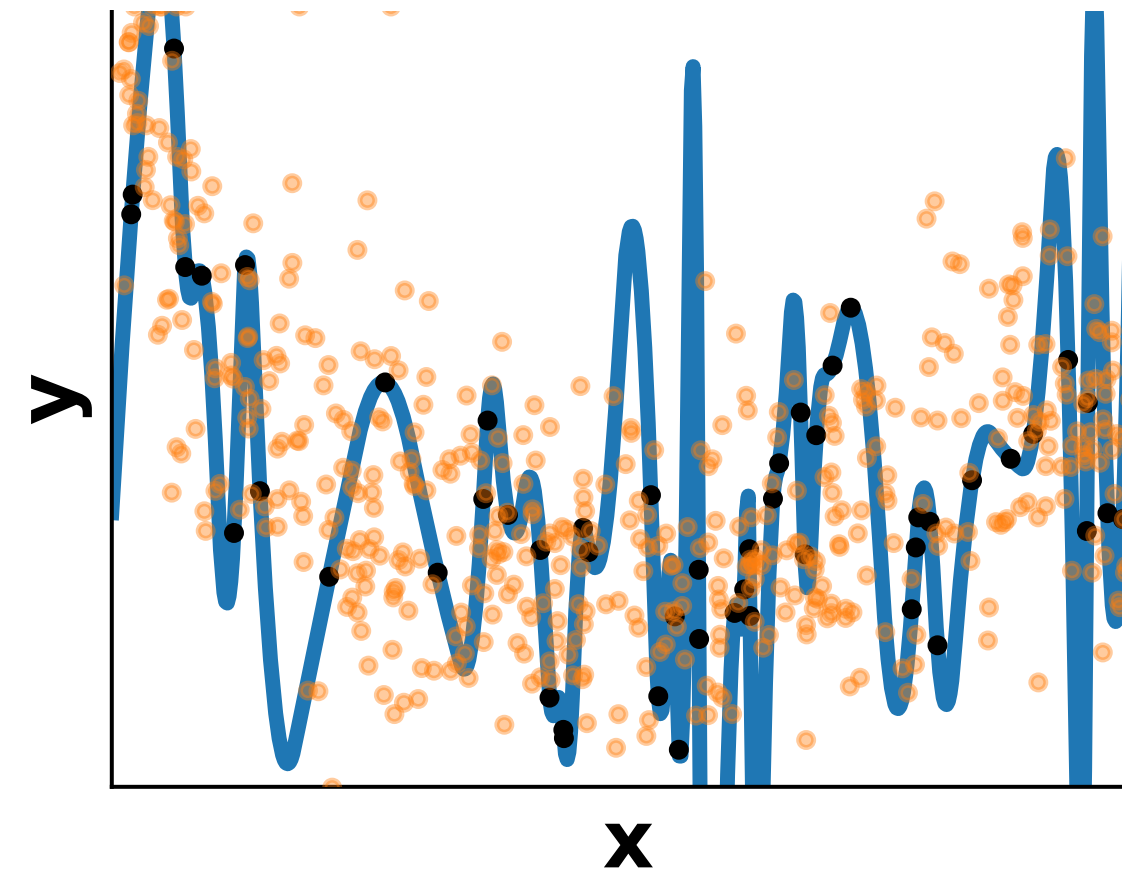
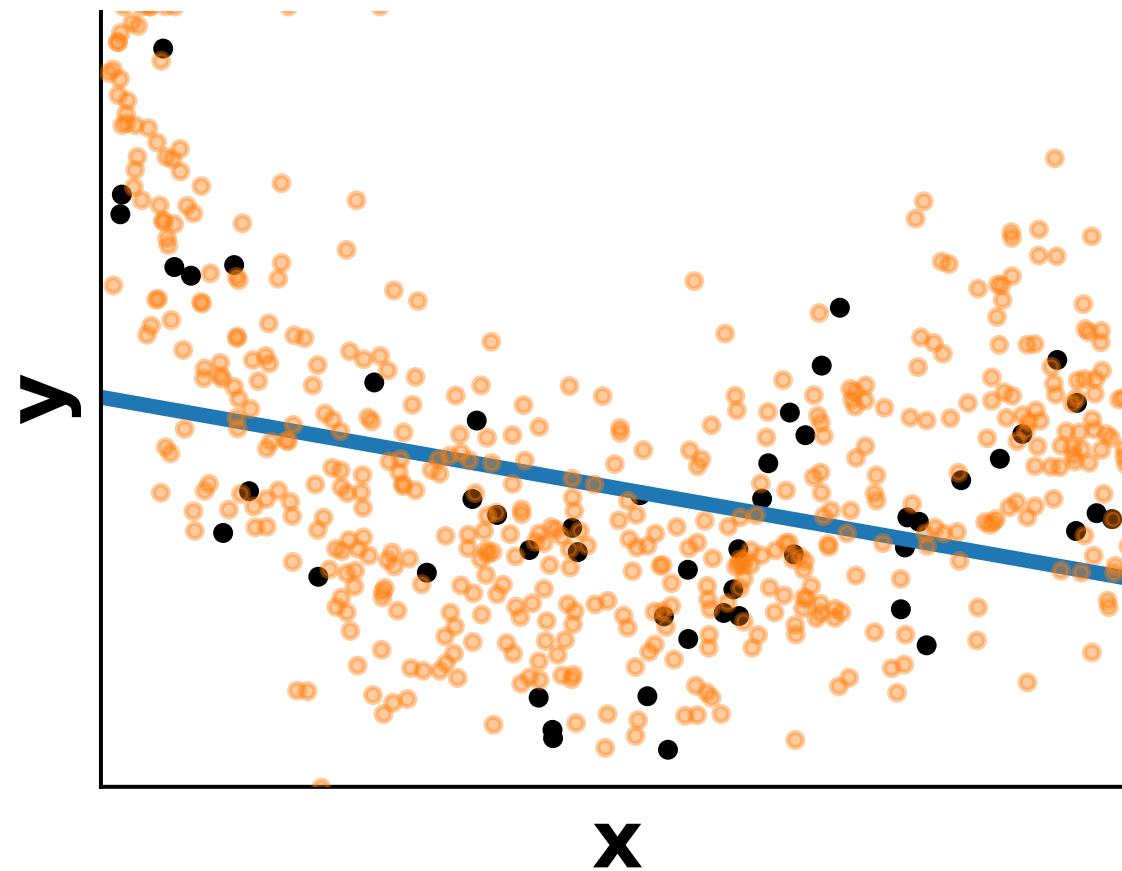
Which data fit do you prefer?



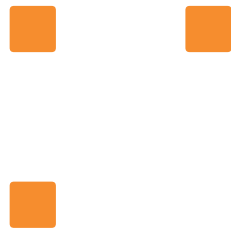
On new data



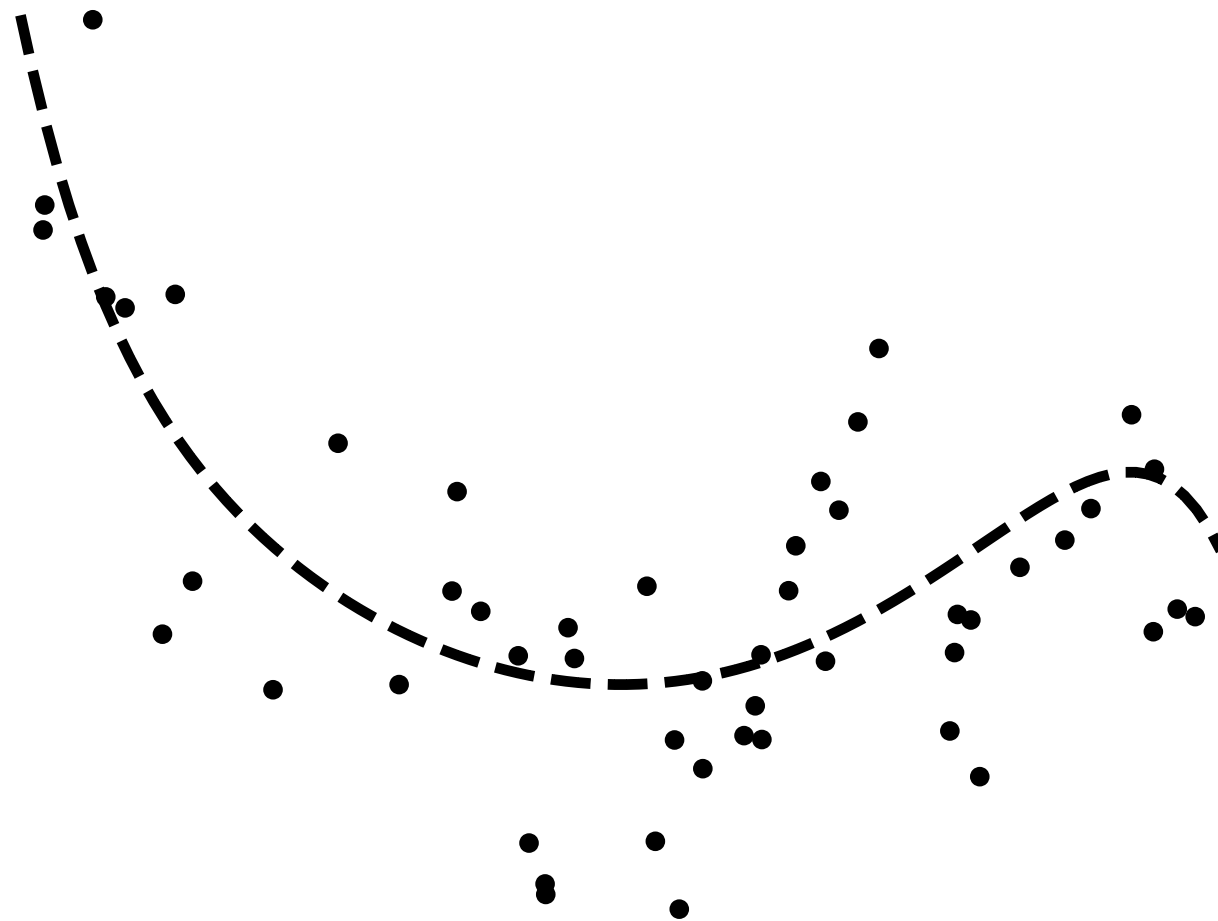
Which data fit do you prefer?



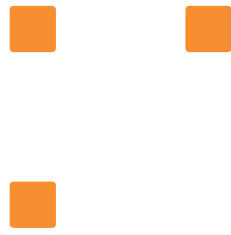
A bit more complex example



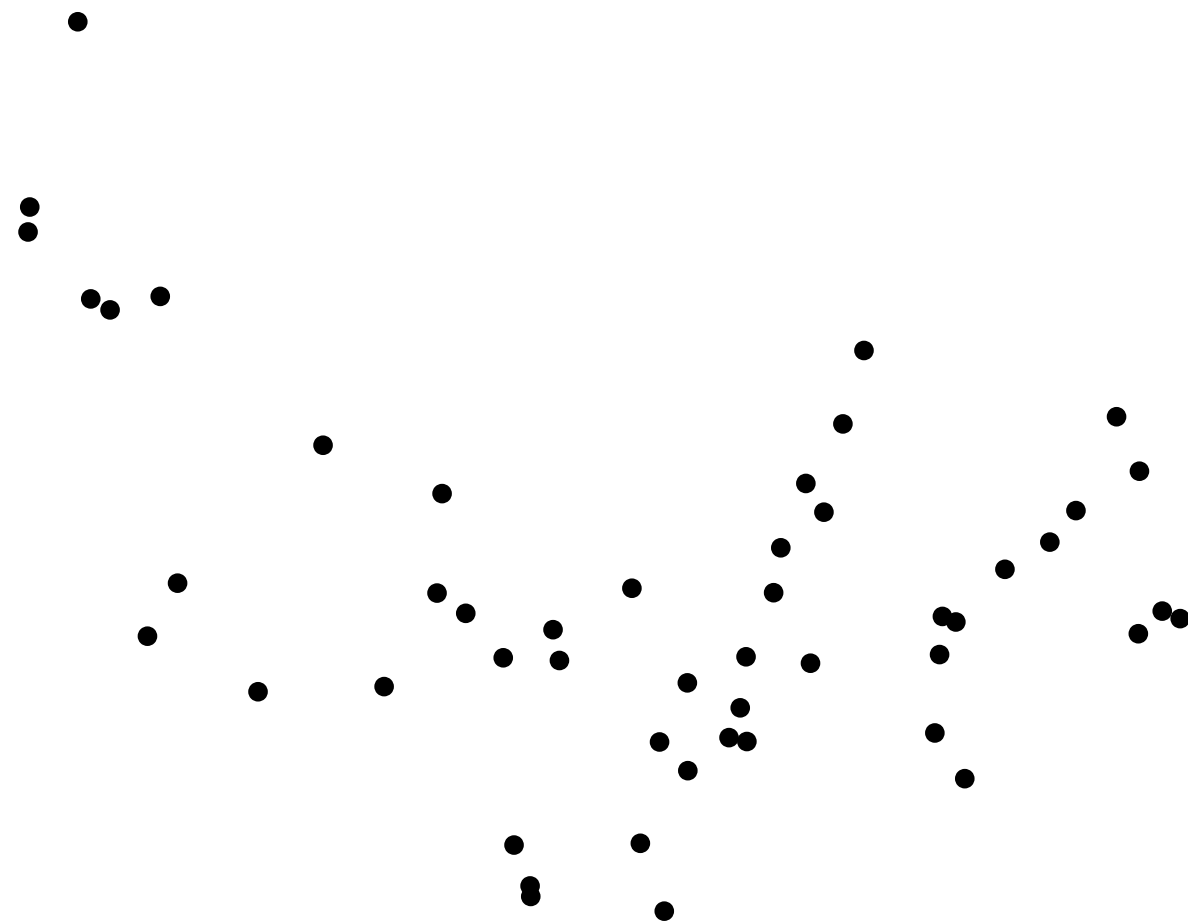
Varying model complexity



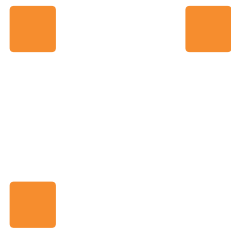
- Data generated with a 9th-degree polynomial + some noise



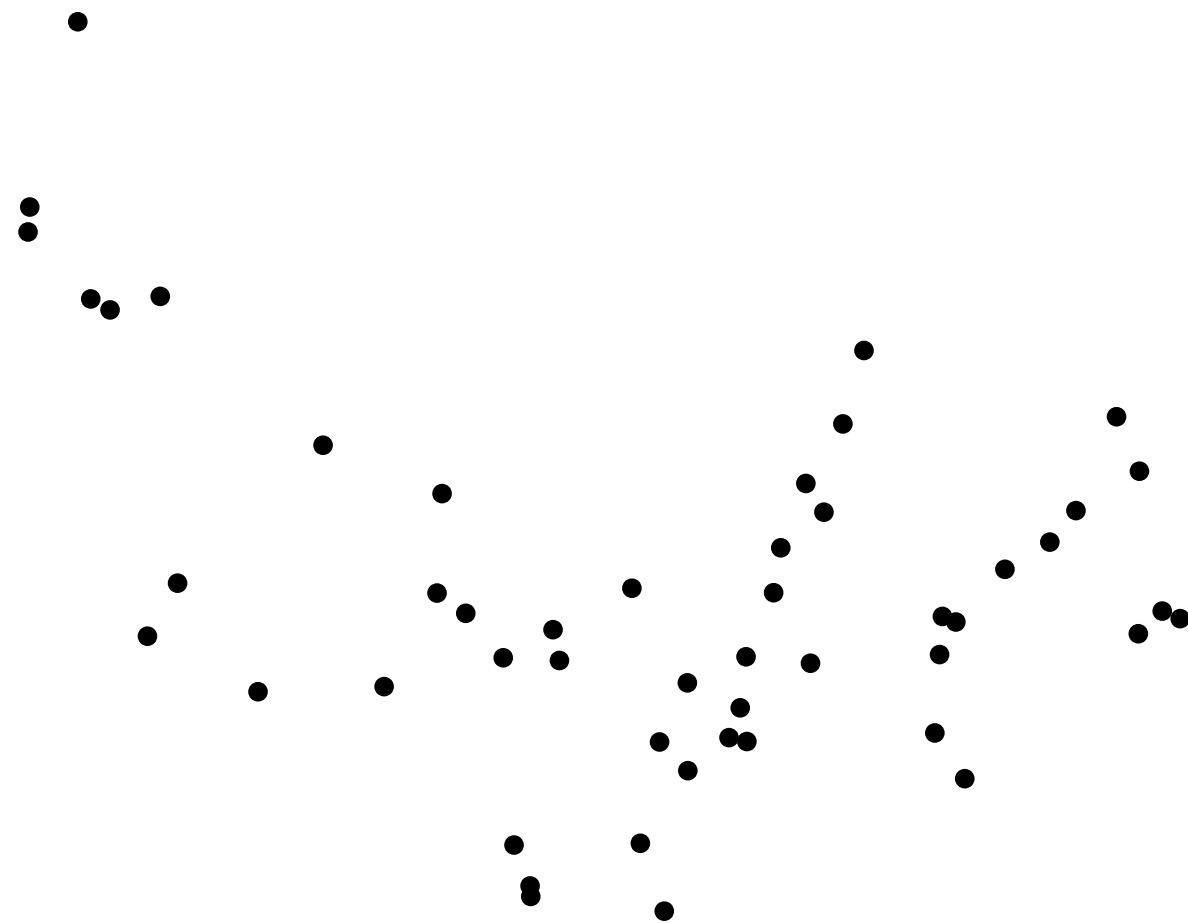
Varying model complexity



- Data generated with a 9th-degree polynomial + some noise
- This process is unknown

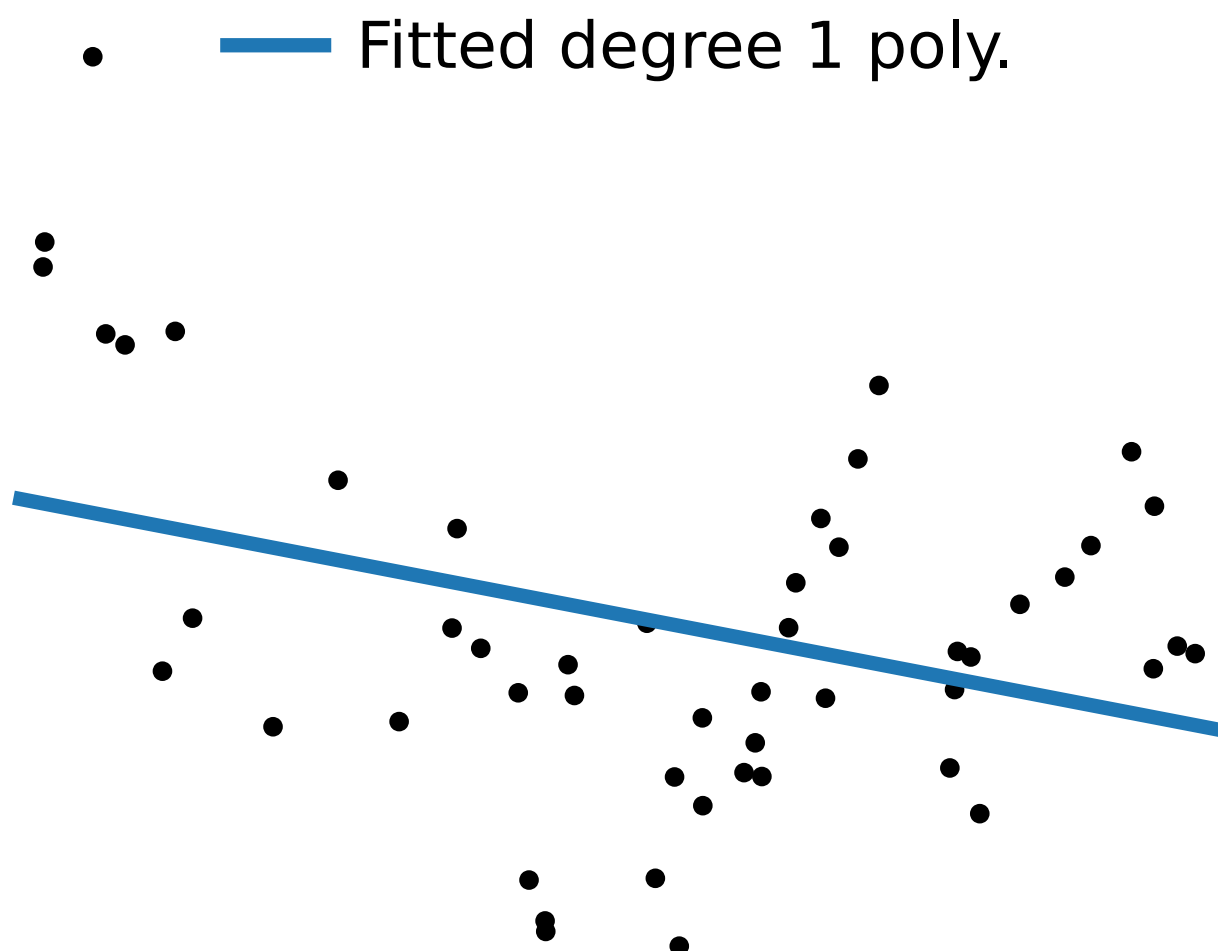


Varying model complexity



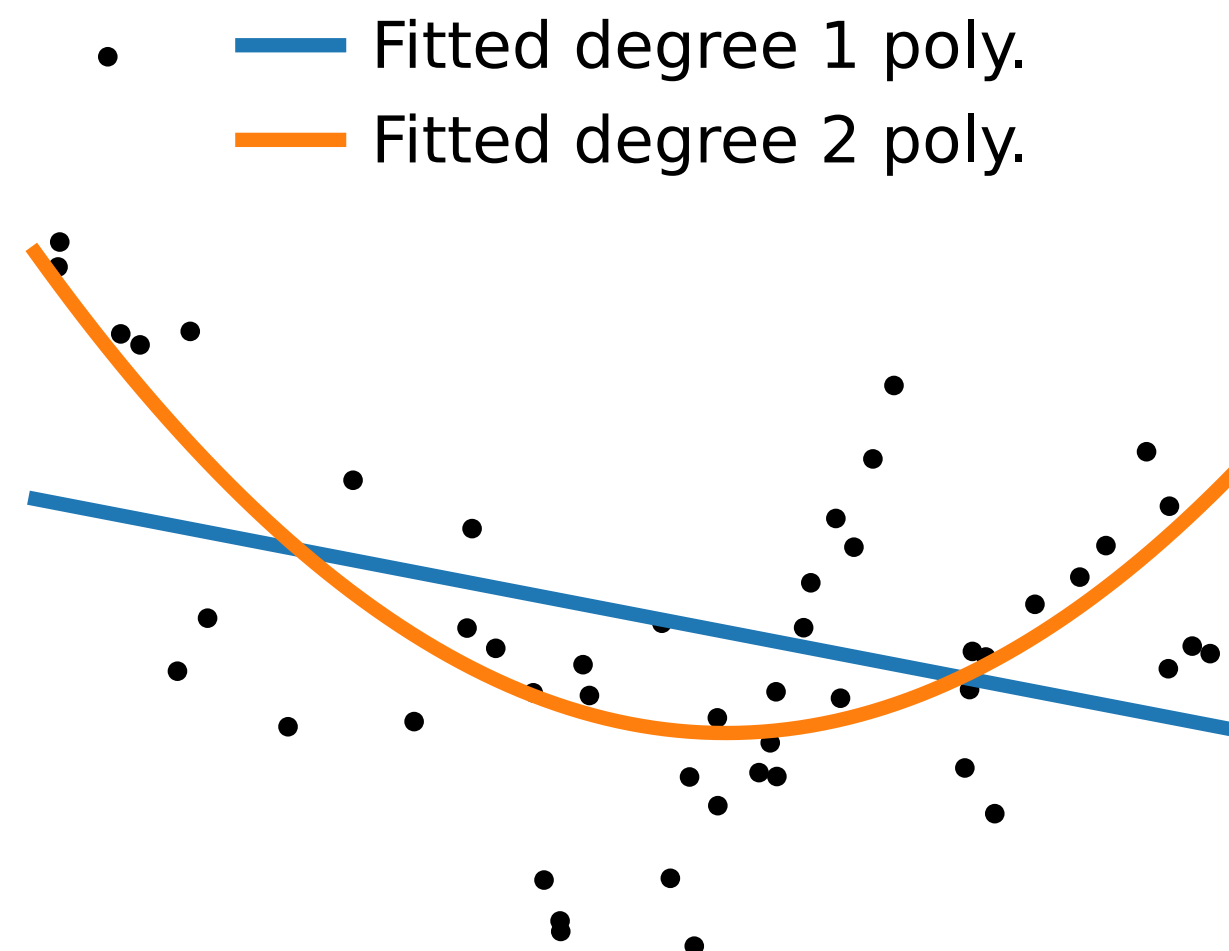
- Data generated with a 9th-degree polynomial + some noise
- This process is unknown
- We can only access the observations

Varying model complexity



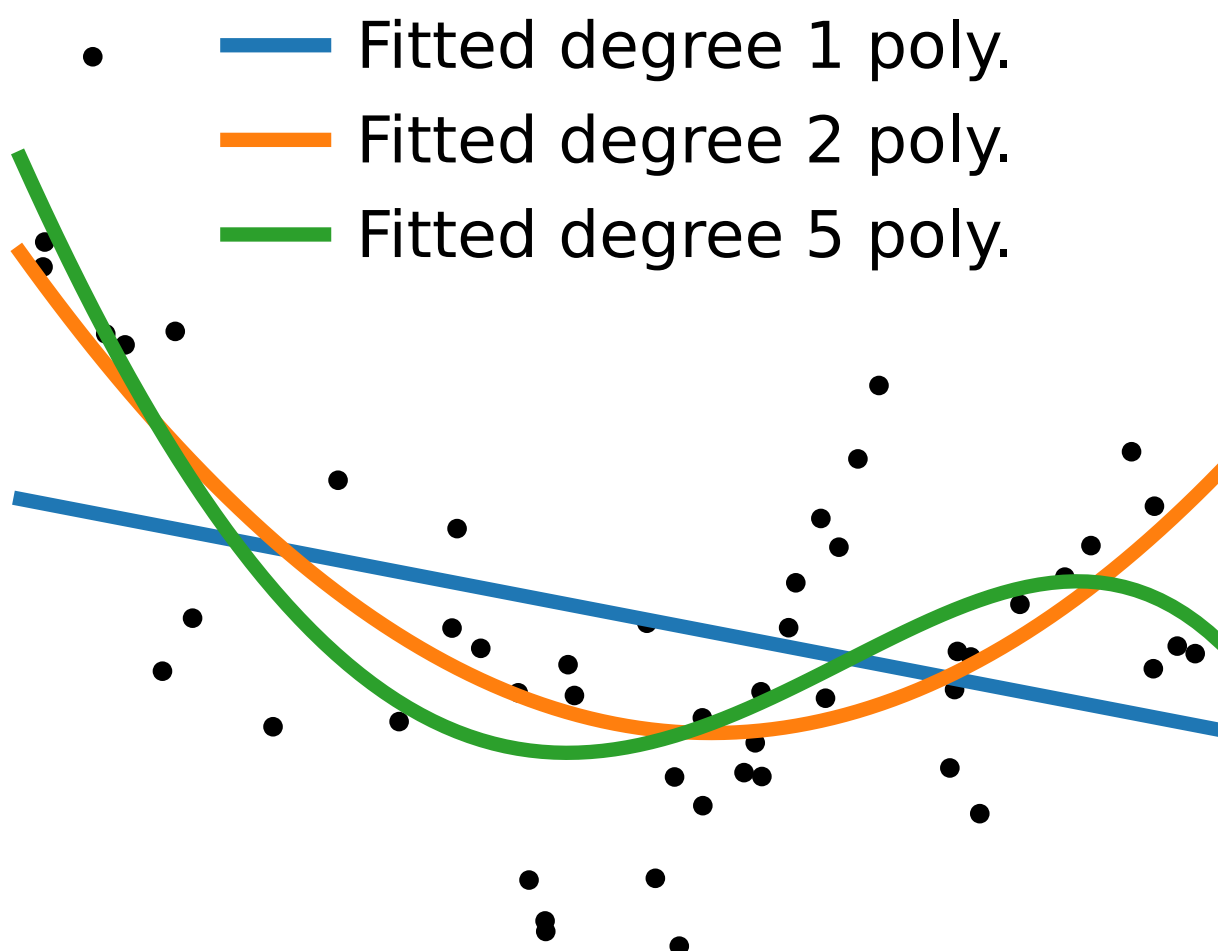
- Data generated with a 9th-degree polynomial + some noise
- This process is unknown
- We can only access the observations
- Fit polynomials of various degrees

Varying model complexity



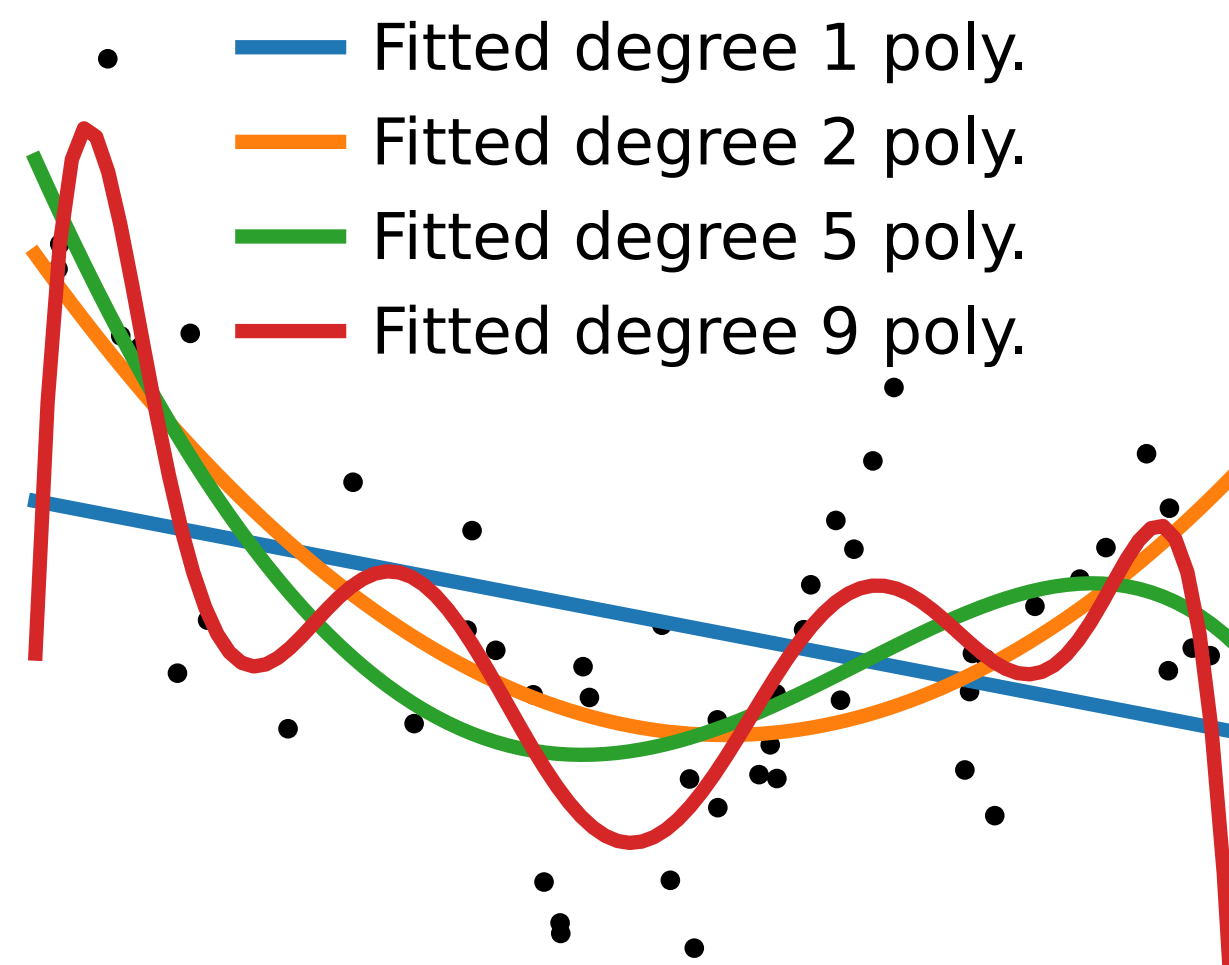
- Data generated with a 9th-degree polynomial + some noise
- This process is unknown
- We can only access the observations
- Fit polynomials of various degrees

Varying model complexity



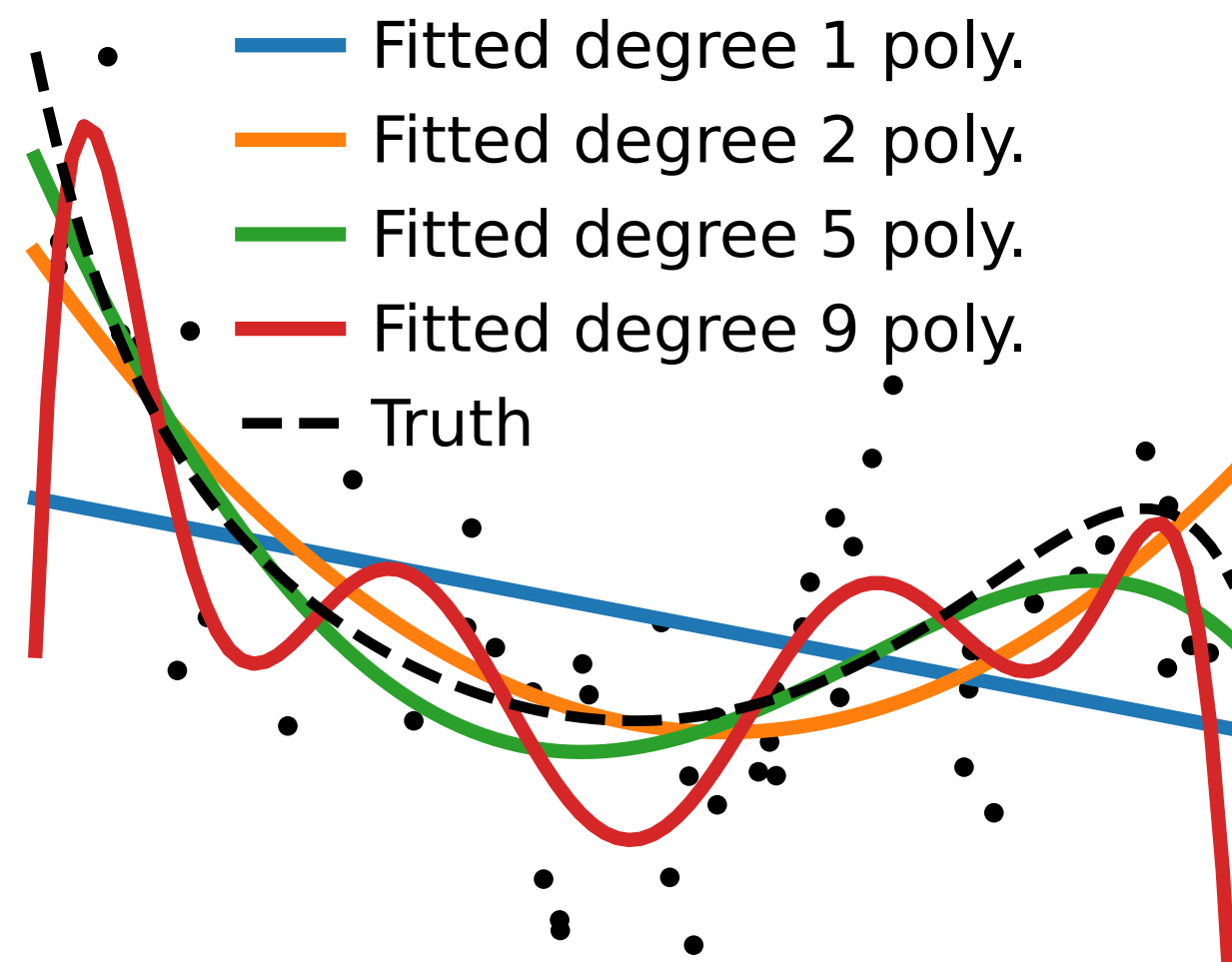
- Data generated with a 9th-degree polynomial + some noise
- This process is unknown
- We can only access the observations
- Fit polynomials of various degrees

Varying model complexity



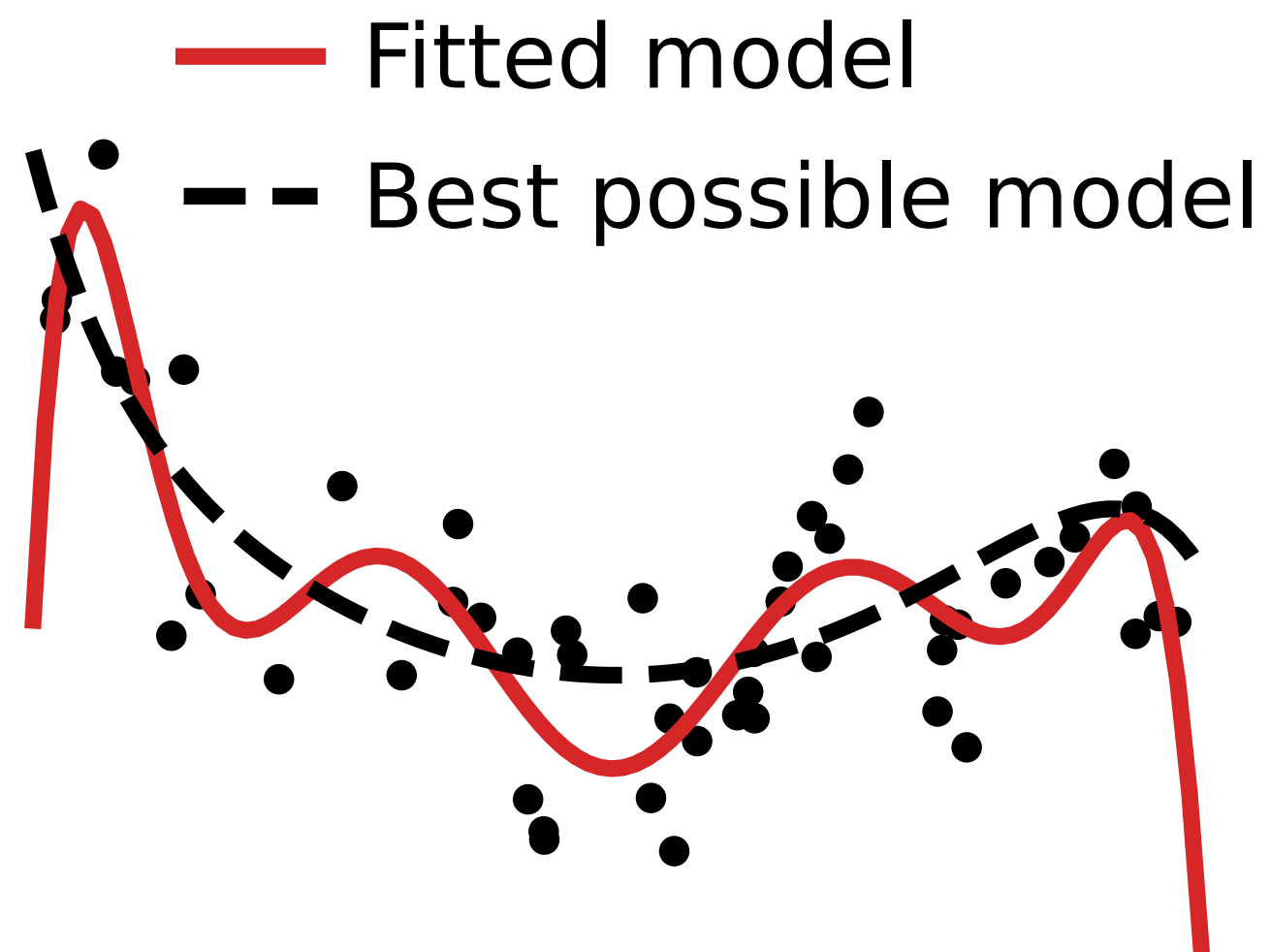
- Data generated with a 9th-degree polynomial + some noise
- This process is unknown
- We can only access the observations
- Fit polynomials of various degrees

Varying model complexity



- Data generated with a 9th-degree polynomial + some noise
- This process is unknown
- We can only access the observations
- Fit polynomials of various degrees

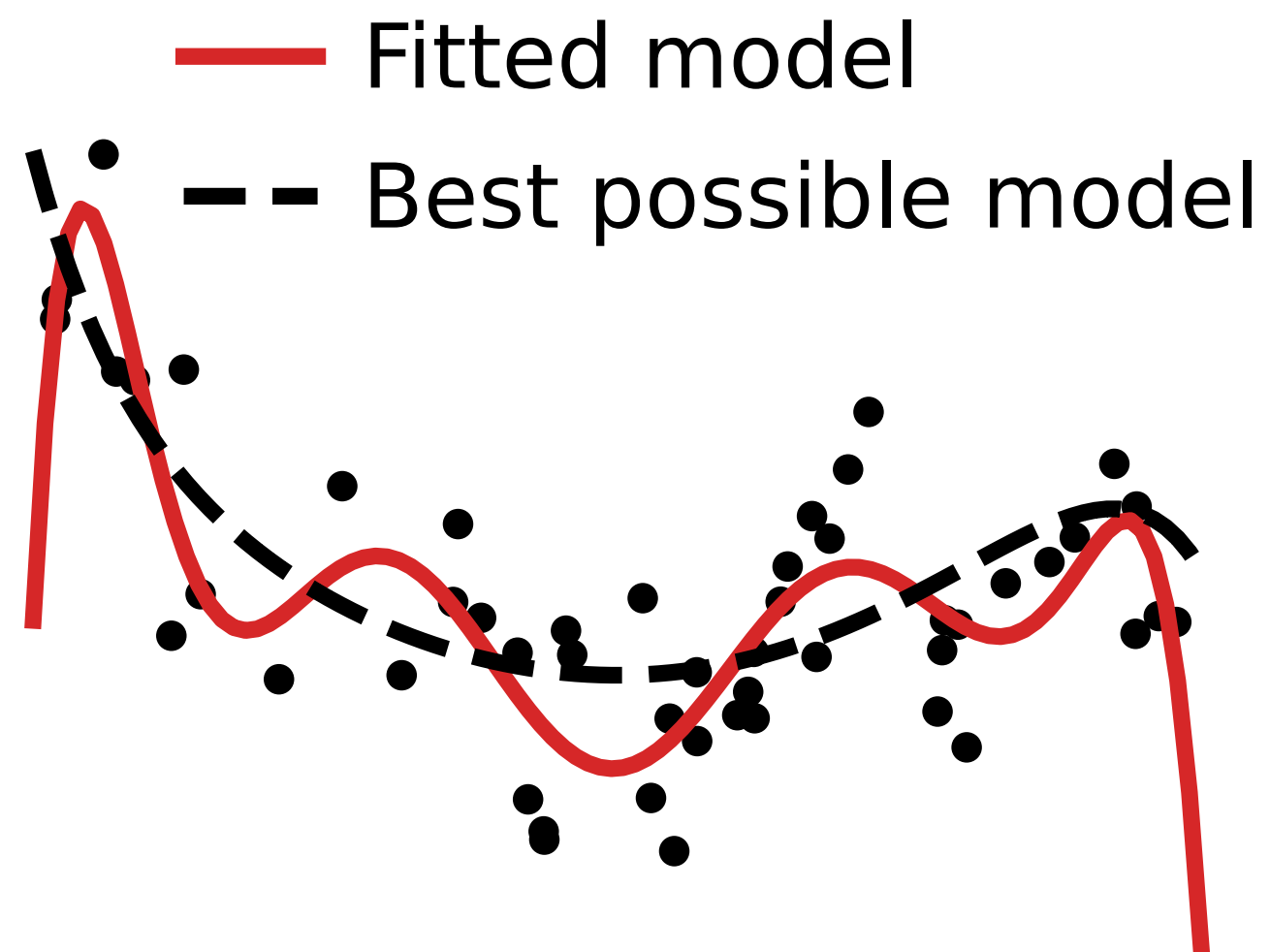
Overfit: model too complex



Model too complex for the data:

- Its best possible fit would approximate well the generative process
- However, its flexibility captures noise

Overfit: model too complex

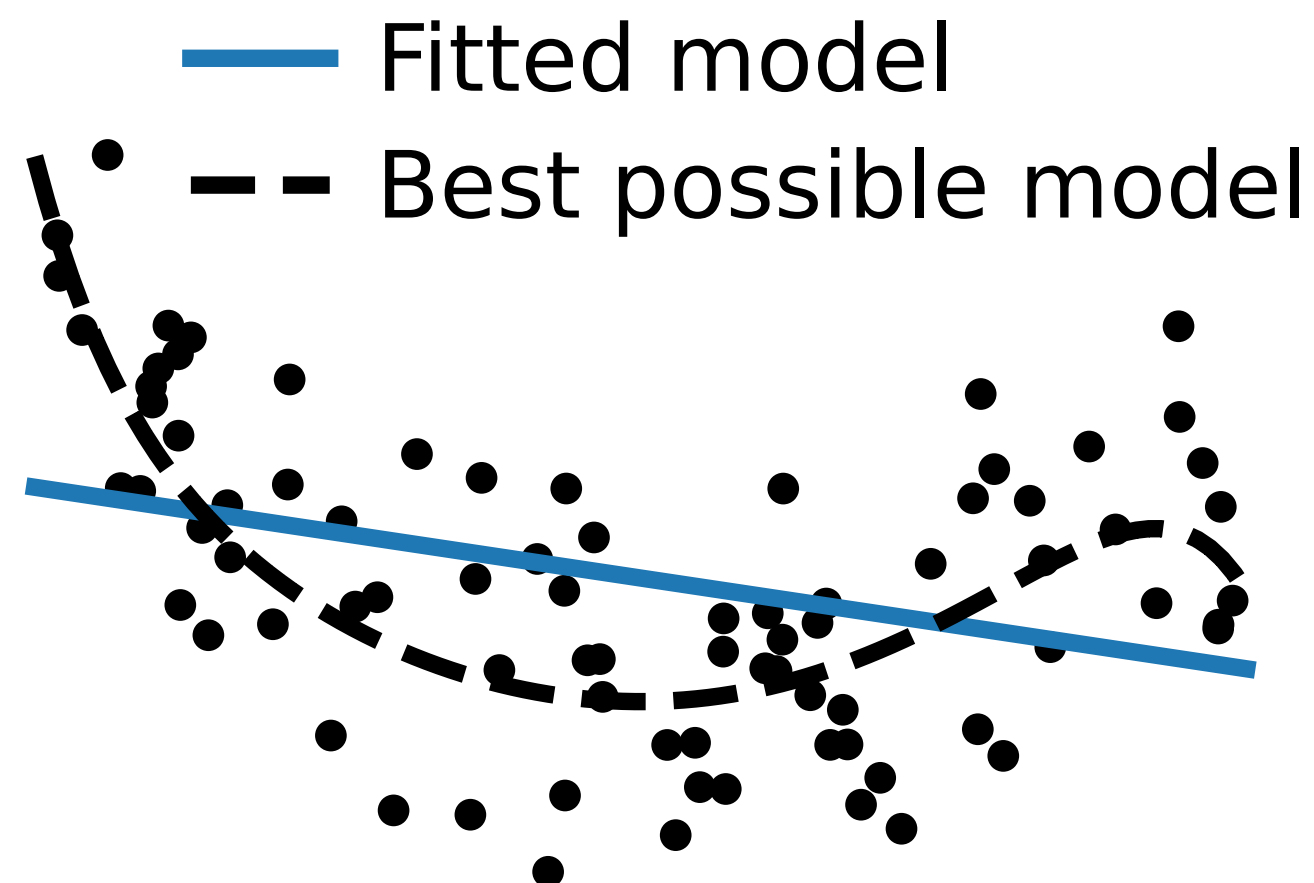


Model too complex for the data:

- Its best possible fit would approximate well the generative process
- However, its flexibility captures noise

Not enough data Too much noise

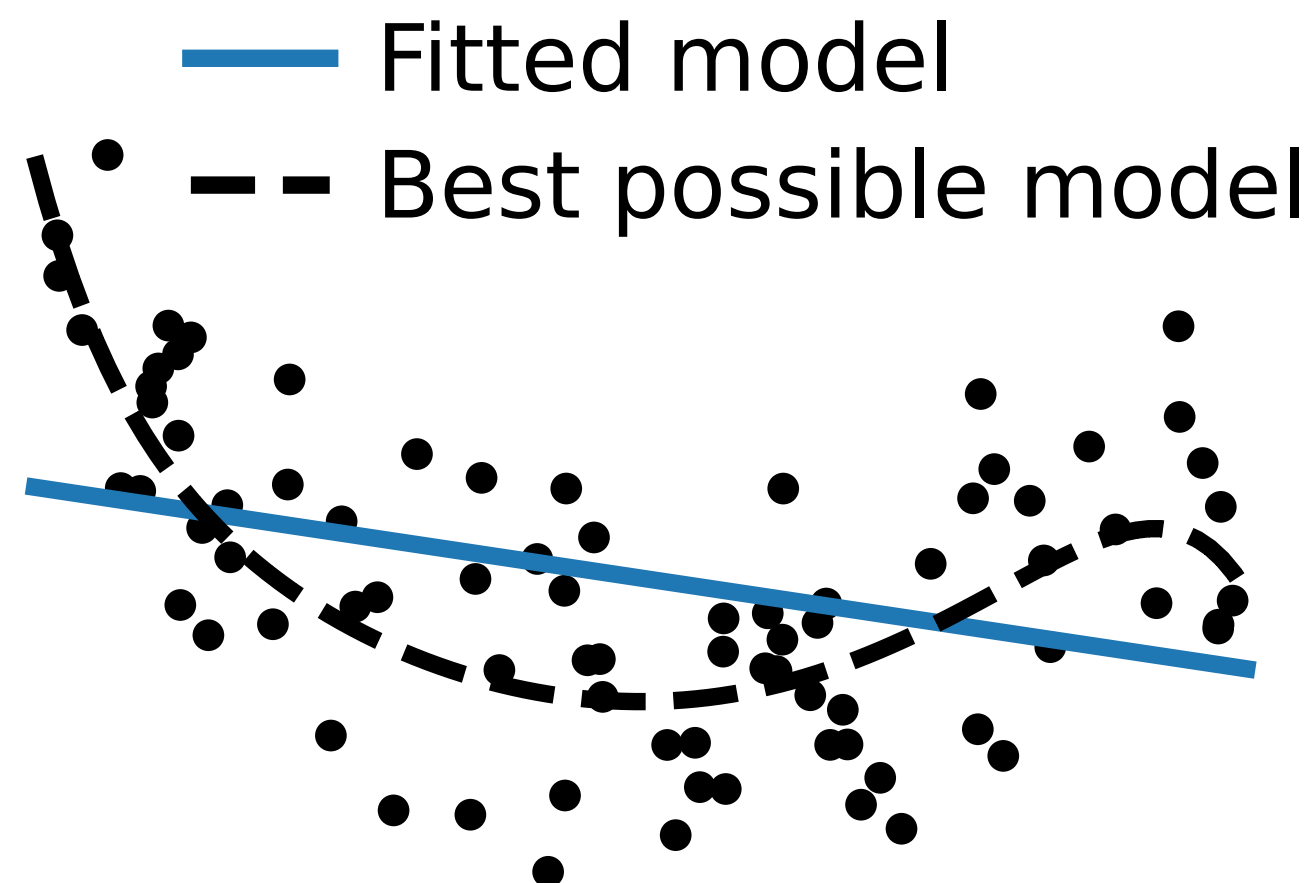
Underfit: model too simple



Model too simple for the data:

- Its best does not approximate well the generative process
- Yet it captures little to no noise

Underfit: model too simple

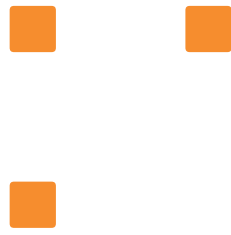


Model too simple for the data:

- Its best does not approximate well the generative process
- Yet it captures little to no noise

Plenty of data

Low noise



Main takeaways

Models too complex for the data overfit:

- They explain too well the data that they have seen
- They do not generalize

Models too simple for the data underfit;

- They capture no noise
- They are limited by their expressivity

How to find the right trade-off?