IQ test

- extend the sequence: 0, 1, 1, 2, 3, 5, 8, . . .
- extend the sequence: 0, 1, 1, 2, 5, 9, 18, . . .
- more interesting is to find a systematic solution
- the key ingredient is rank deficiency of a matrix

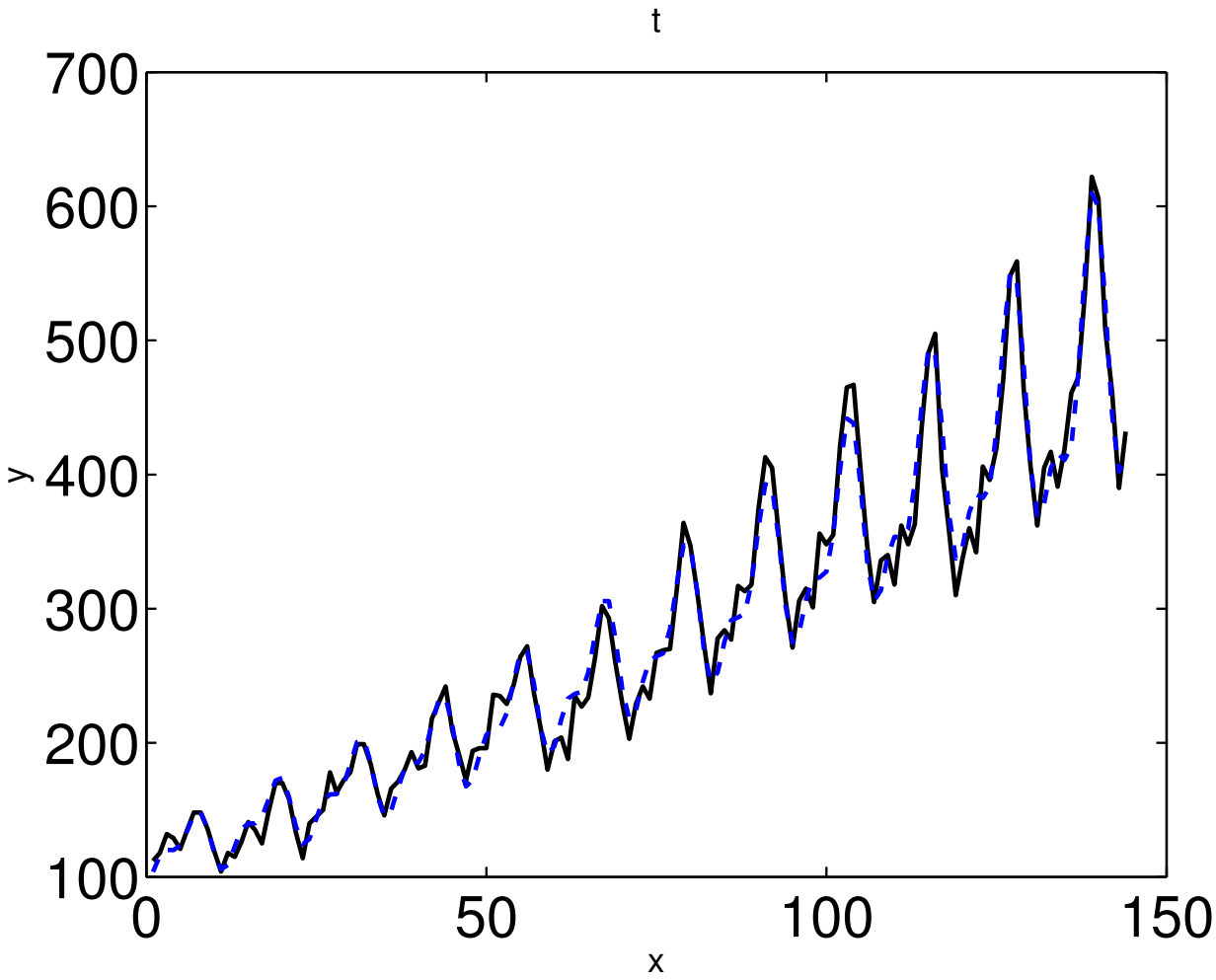
  "Behind every data modeling problem there is a (hidden) low-rank approximation problem: the model imposes relations on the data which render a matrix constructed from exact data rank deficient."

Time series interpolation

- from extrapolation to interpolation
- data: classic Box & Jenkins airline data
  monthly airline passenger numbers 1949–1960
- aim: estimate missing values
  - missing values in "the future": extrapolation
  - other missing values: interpolation
  - take into account the time series nature of the data

Autonomous LTI model

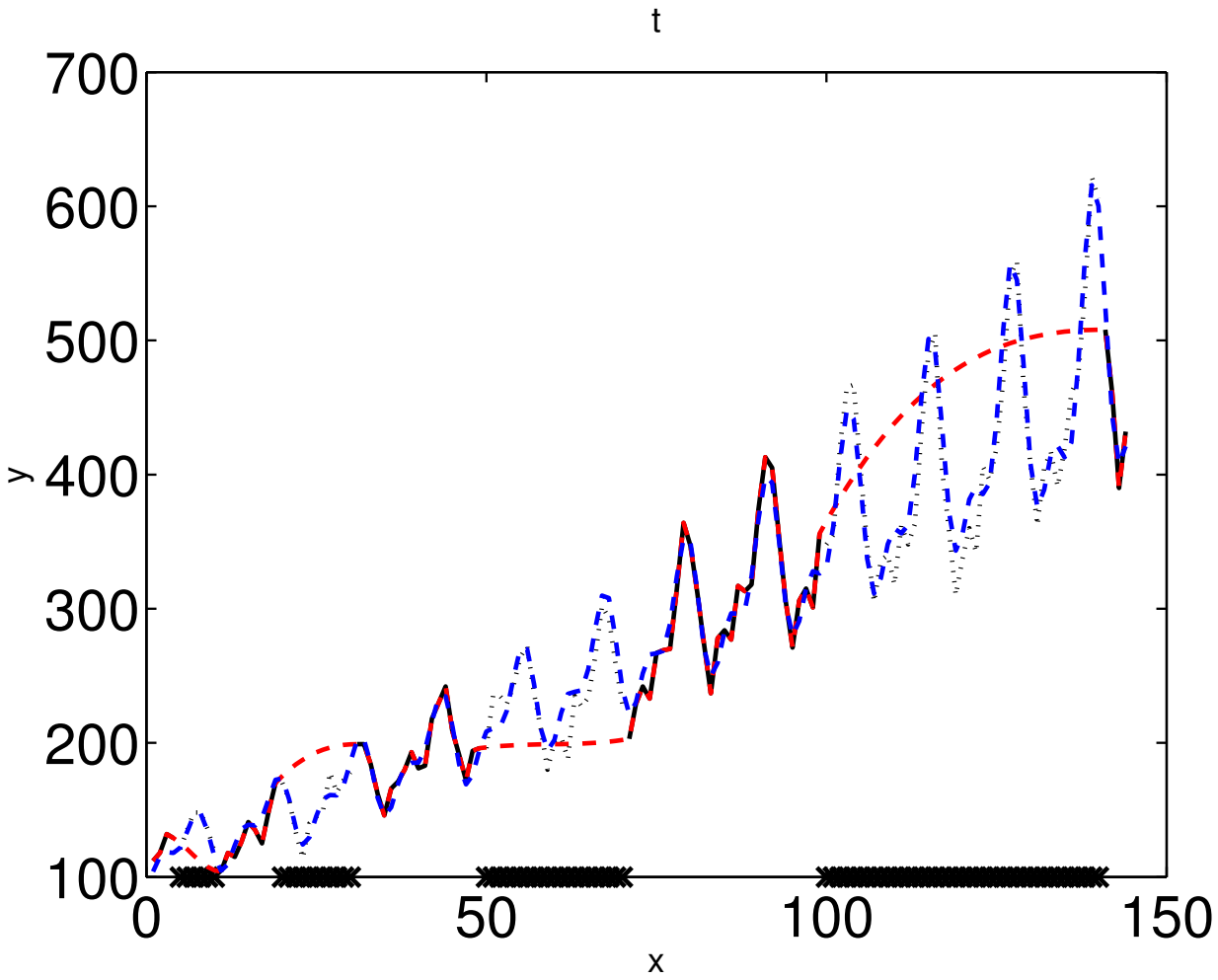
- using all 144 data points to identify a model



- solid line — data, dashed — fit by 6th order model

Missing data estimation

- `[5:10 20:30 50:70 100:140]` are missing



- piecewise cubic interpolation, 6th order LTI model

Modeling as data compression

- the model is a concise representation of the data
- exact model $\leftrightarrow$ lossless compression (*e.g.*, `zip`)
- approximate model $\leftrightarrow$ lossy compression (*e.g.*, `mp3`)

Example: compression of a random vector

- data: $1 \times n$ vector, generated by `randn`
- compression in `mat` format

	length n	1	223	334	556	667	1000
1.	original size	8	1784	2672	4448	5336	8000
2.	mat file size	178	1945	2798	4490	5341	7893

Example: low-rank matrix compression

- data: random 100×100 matrix D of rank 5
- stored in four different ways

	representation	size
1.	all elements of D	80000
2.	D in `mat` format	75882
3.	all elements of P and L	8024
4.	P and L in `mat` format	7767

- in 2 and 4, we compute a rank revealing factorization

$$D = PL$$

- can we do better than storing P and L (compressed)?

Example: trajectory of an LTI system

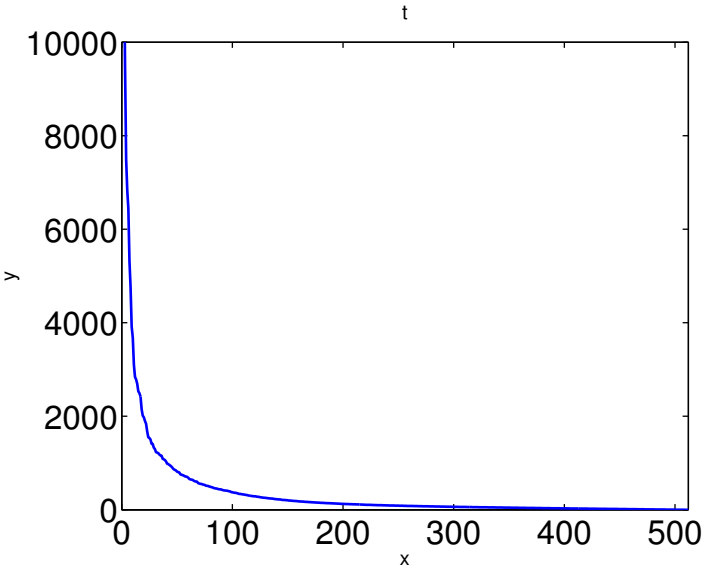
- data: impulse response of a random 3rd order system
- stored in four different ways

	representation	size
1.	impulse response h	192
2.	h in `mat` format	377
3.	model parameters θ	56
4.	θ in `mat` format	233

- in 3 and 4, we have parameterized the system

Low-rank approximation of images

- an image is a matrix of gray values (integers 0–255)
- typical singular values plot:



- $\Longrightarrow$ an image can be approximate by lower rank
- the basis of many methods for image processing
- note that SVD does not respect the 0–255 bounds

Original 512×512 image

Rank 100 approximation

Rank 80 approximation

Rank 60 approximation

Rank 40 approximation