Towards Big Data Cybernetics: examples of useful algorithms

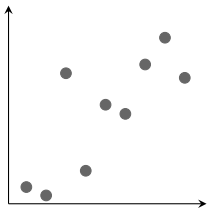
NTNU, Department of Engineering Cybernetics

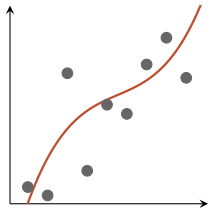
Nov. 27, 2019

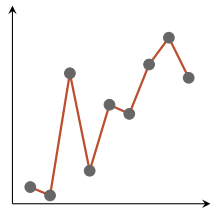
Principal Component Analysis – what and why through an example

(nasa.org)

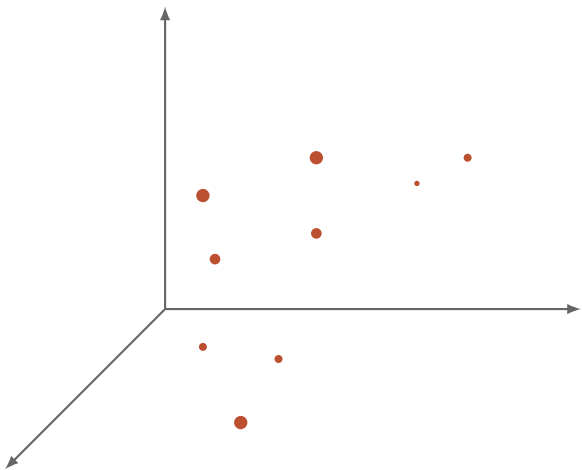
To compress = extract all the information – but not more!



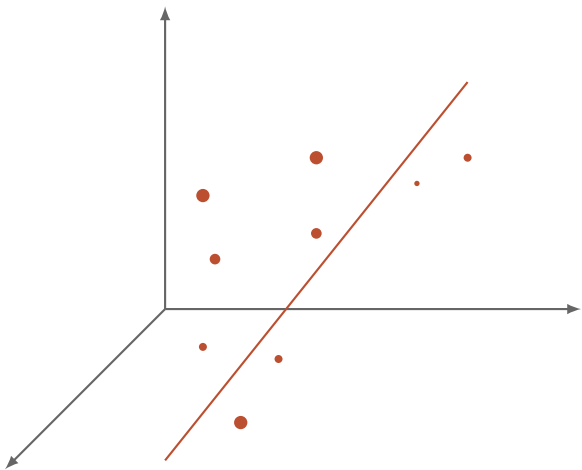




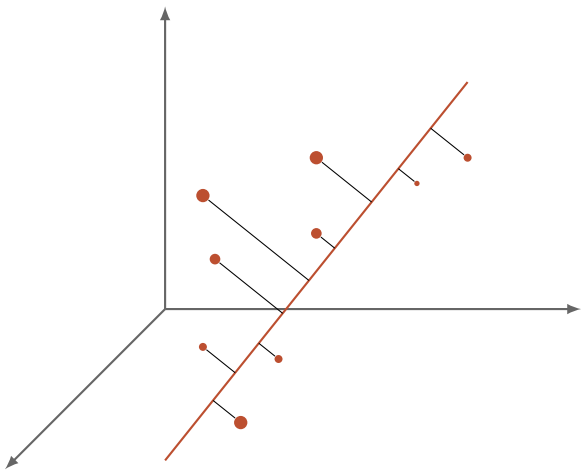
Principal Component Analysis – a graphical explanation



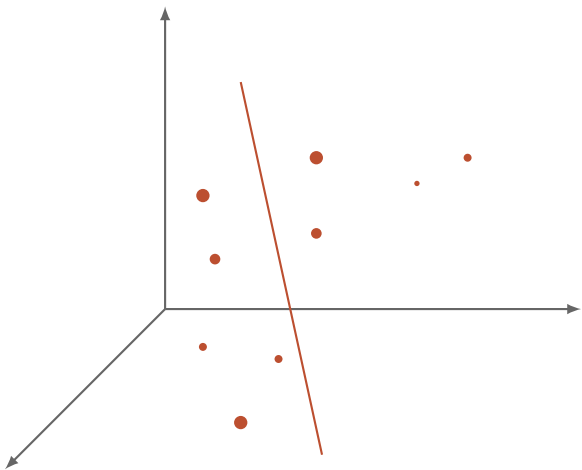
Principal Component Analysis – a graphical explanation



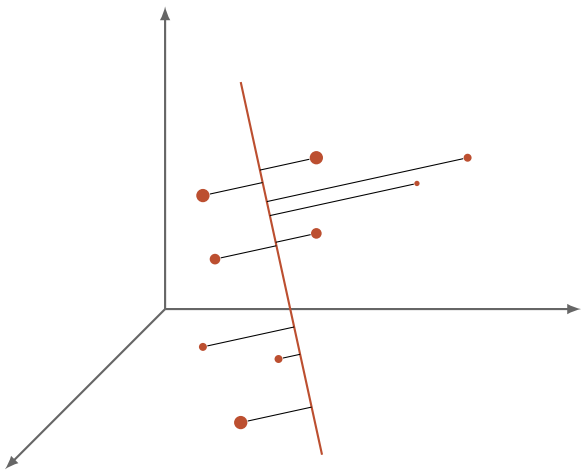
Principal Component Analysis – a graphical explanation



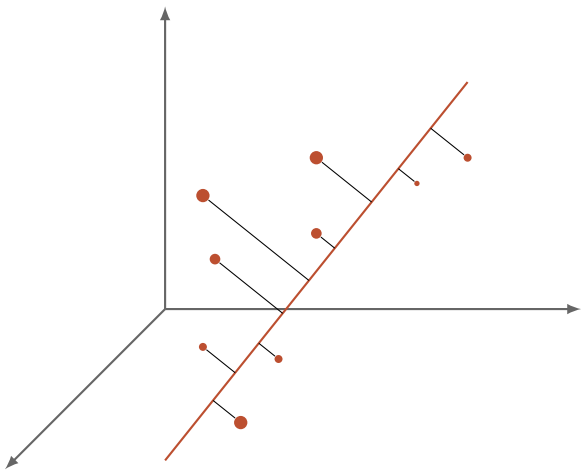
Principal Component Analysis – a graphical explanation



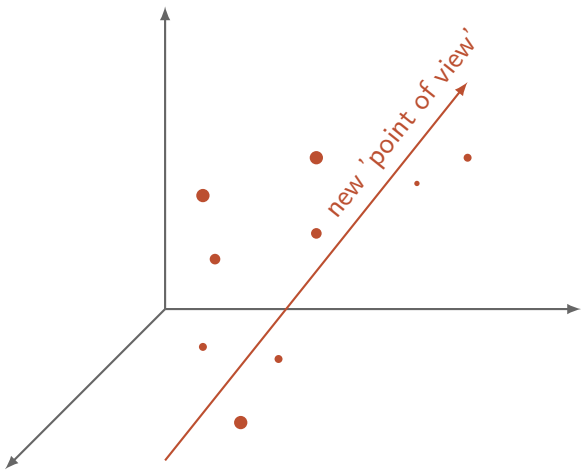
Principal Component Analysis – a graphical explanation



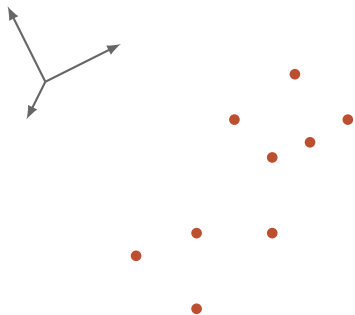
Principal Component Analysis – a graphical explanation



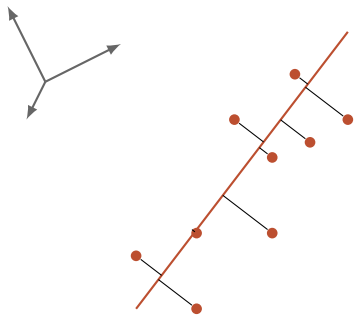
Principal Component Analysis – a graphical explanation



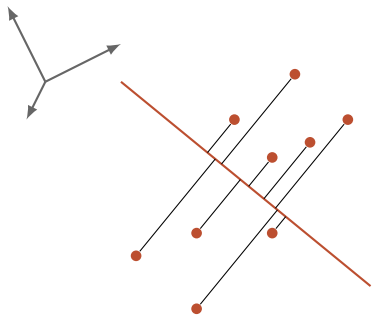
Principal Component Analysis – a graphical explanation



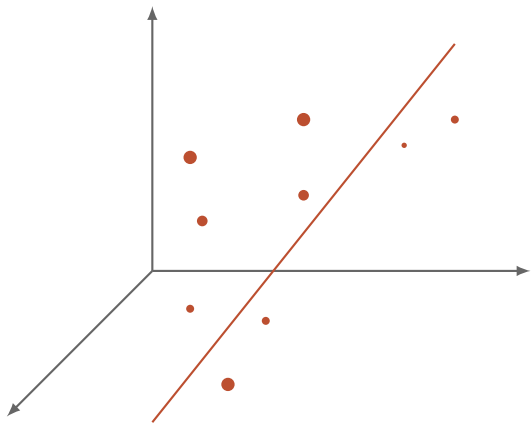
Principal Component Analysis – a graphical explanation



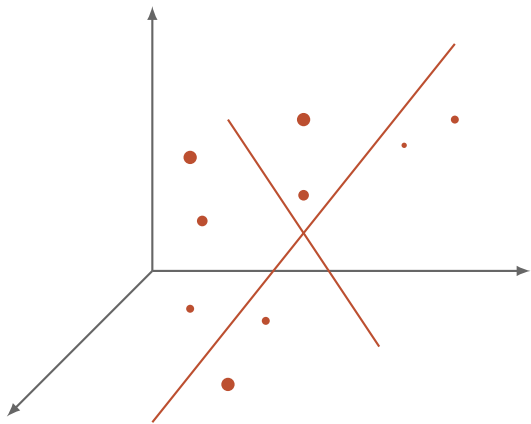
Principal Component Analysis – a graphical explanation



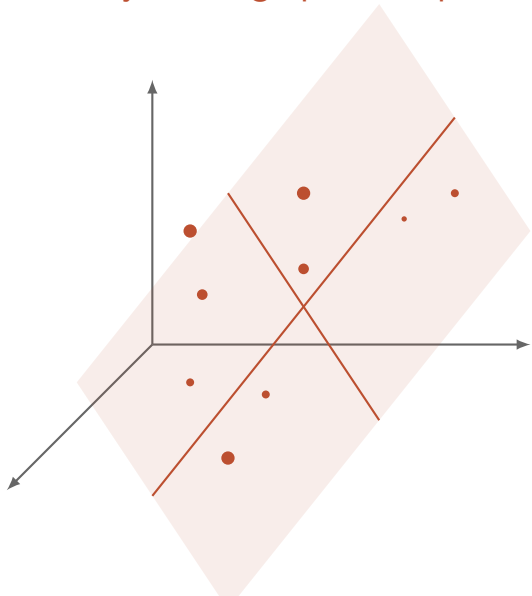
Principal Component Analysis – a graphical explanation



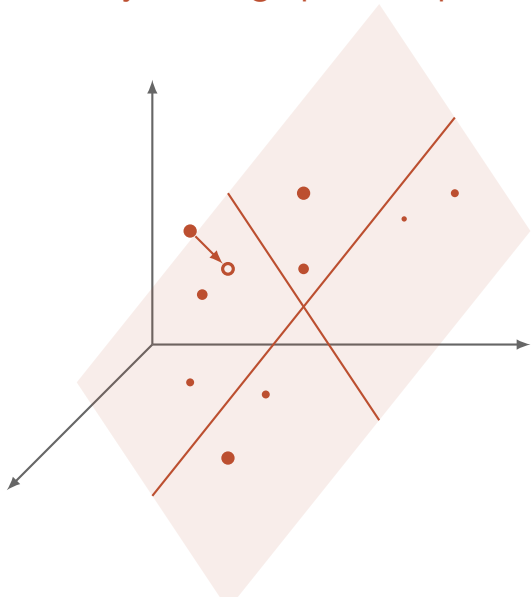
Principal Component Analysis – a graphical explanation



Principal Component Analysis – a graphical explanation



Principal Component Analysis – a graphical explanation



Principal Component Analysis

What, why, and its connections with Big Data Cybernetics

- linear transformation that most accurately reconstructs the training data in a geometric distance sense
- answers the need for simplification / compression
- very fast, & can include missing data and nonlinearities

10 seconds of pause

PLS – what and why through an example

x_1	x_2	$\ldots$	x_n	y_1	$\ldots$	y_m

PLS – what and why through an example

x_1	x_2	$\ldots$	x_n	y_1	$\ldots$	y_m
0.87	11.2	$\ldots$	−3.93	*ok*	$\ldots$	44
0.67	14.4	$\ldots$	−2.89	*ok*	$\ldots$	39
$\vdots$	$\vdots$		$\vdots$	$\vdots$		$\vdots$
0.91	12.7	$\ldots$	−1.65	*not ok*	$\ldots$	15

PLS – what and why through an example

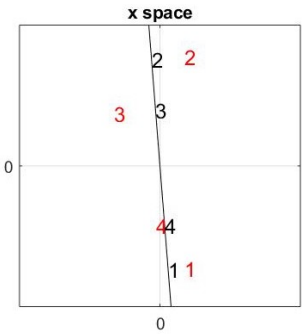
x_1	x_2	...	x_n	y_1	...	y_m
0.87	11.2	...	−3.93	*ok*	...	44
0.67	14.4	...	−2.89	*ok*	...	39
⋮	⋮		⋮	⋮		⋮
0.91	12.7	...	−1.65	*not ok*	...	15

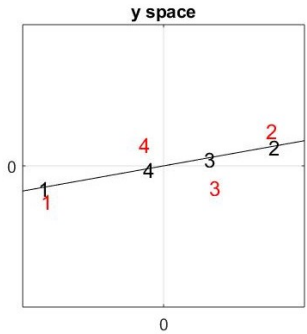
0.74	13.1	...	−3.51	?	...	?
0.72	13.6	...	−3.43	?	...	?
⋮	⋮		⋮	⋮		⋮
0.88	12.5	...	−3.65	?	...	?

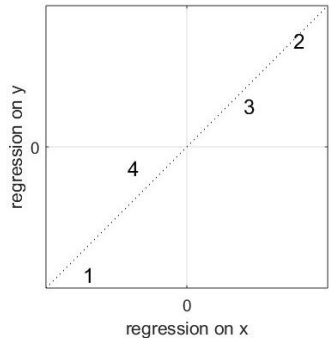
PLS – what and why through an example

http://folk.ntnu.no/damianov/Teaching/TK8117/Videos/PLSanimation.mp4

PLS – what and why through an example







PLS – connections with Big Data Cybernetics

- can handle missing data
- answers the need for simplification / compression
- extremely fast & can handle huge amounts of data
- is interpretable