Summary

Lag features

Motivation for lag features

1. Want to predict future values of the target.

2. Past values of the target are likely to be predictive.

3. Past values of a feature could also be predictive (e.g., the sales on a day is related to advertising (ad) spend on prior days).

target ↓ feature ↓

Date	Sales	Ad spend
2020-02-12	23	100
2020-02-13	30	120
2020-02-14	35	90
2020-02-15	30	80
2020-02-16	?	100

Lag of the target

	target		lag features **from target**		original features	lag features **from original features**	
	Date	**Sales**	**Sales Lag 1**	**Sales Lag 3**	**Ad spend**	**Ad spend Lag 1**	**Ad spend Lag 2**
	2020-02-12	23	NaN	NaN	100	NaN	NaN
	2020-02-13	30	23	NaN	120	100	NaN
	2020-02-14	35	30	NaN	90	120	100
	2020-02-15	30	35	23	80	90	120
	2020-02-16	?	30	30	100	80	90

Lag of features

	target	lag features **from target**		original features	lag features **from original features**	
Date	**Sales**	**Sales Lag 1**	**Sales Lag 3**	**Ad spend**	**Ad spend Lag 1**	**Ad spend Lag 2**
2020-02-12	23	NaN	NaN	100	NaN	NaN
2020-02-13	30	23	NaN	120	100	NaN
2020-02-14	35	30	NaN	90	120	100
2020-02-15	30	35	23	80	90	120
2020-02-16	?	30	30	100	80	90

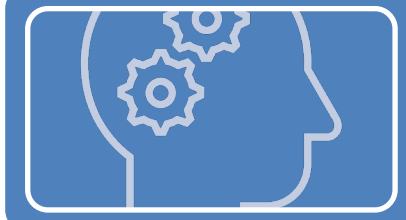
Lag of features

	target	lag features **from target**		original features	lag features **from original features**	
Date	**Sales**	**Sales Lag 1**	**Sales Lag 3**	**Ad spend**	**Ad spend Lag 1**	**Ad spend Lag 2**
2020-02-12	23	NaN	NaN	100	NaN	NaN
2020-02-13	30	23	NaN	120	100	NaN
2020-02-14	35	30	NaN	90	120	100
2020-02-15	30	35	23	80	90	120
2020-02-16	?	30	30	100	80	90

Problem: Which lags to use? How many lag features to create?

How to choose the lags

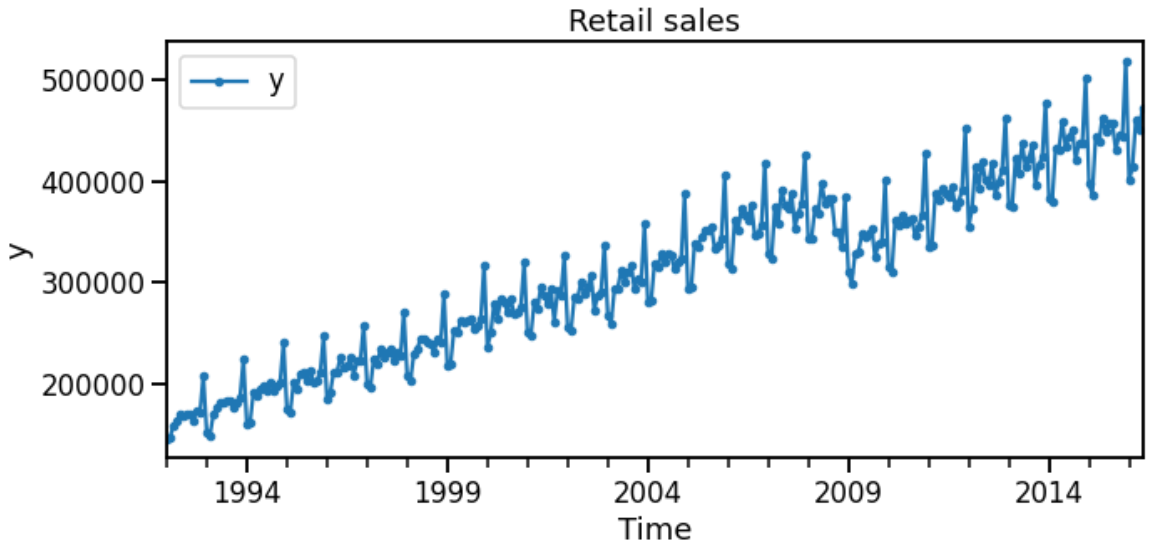
Domain knowledge

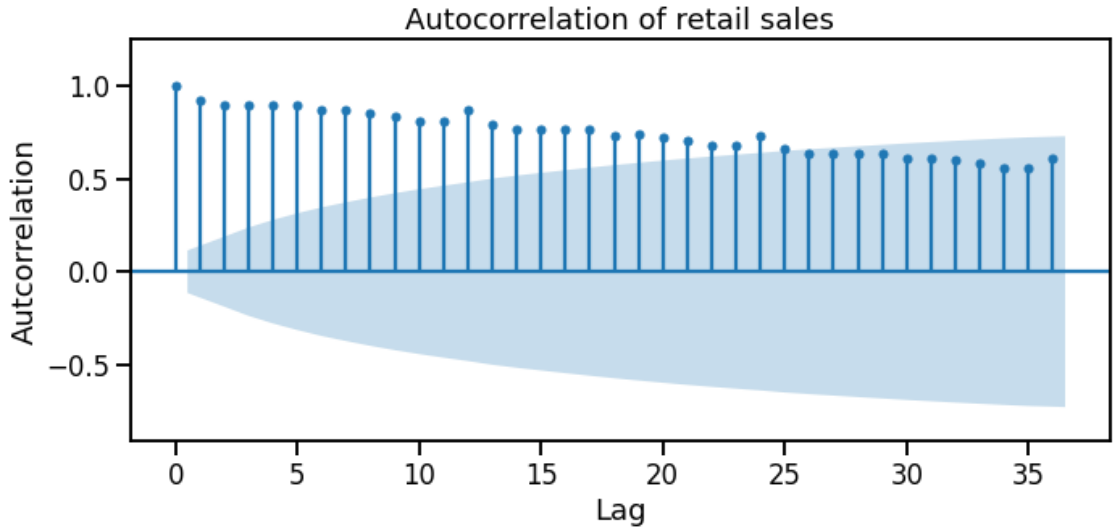


Feature selection and modelling

Time-series correlation methods

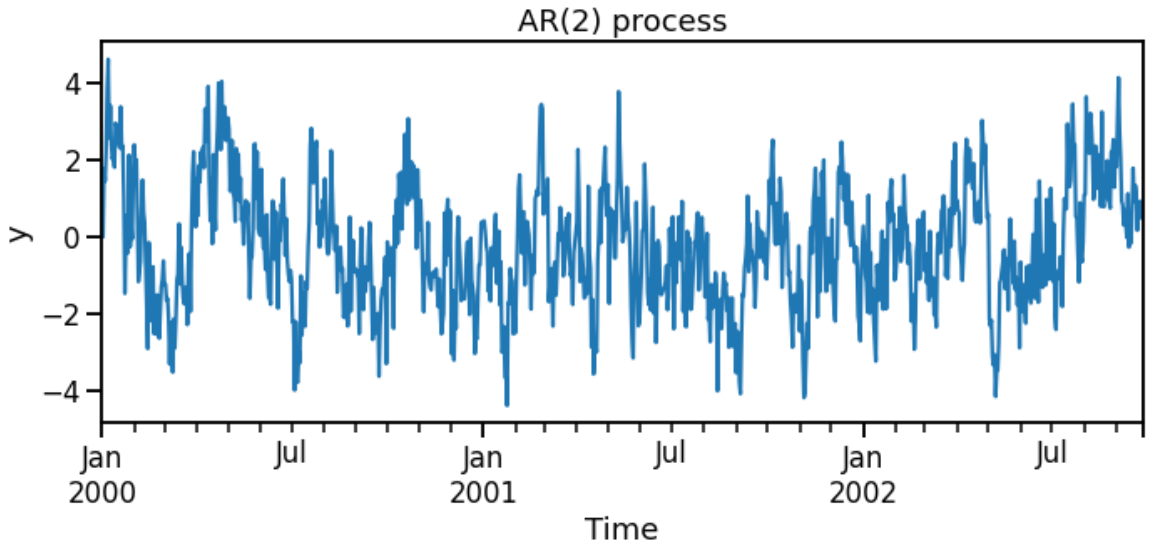
Autocorrelation function (ACF)

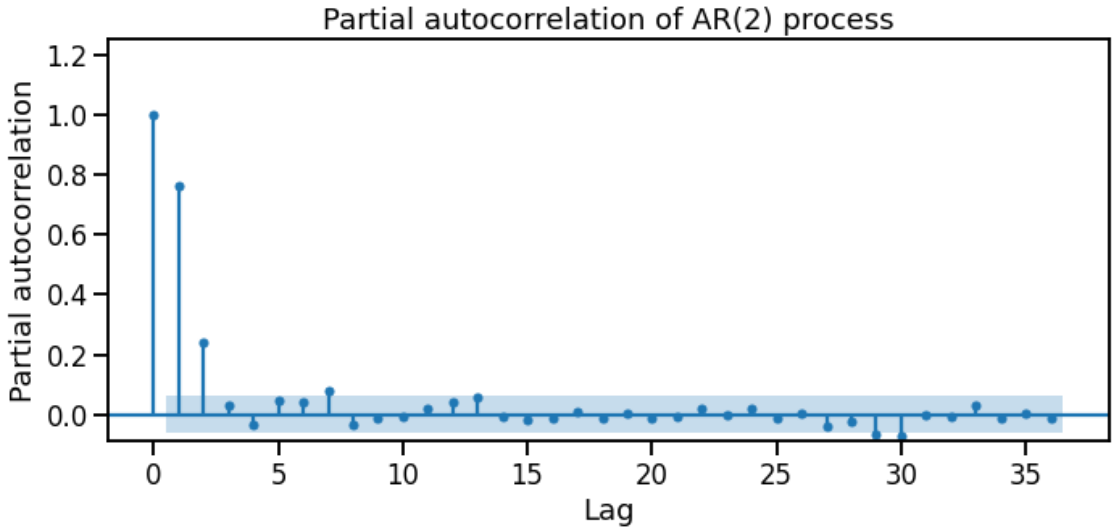




- ACF can identify periodic behaviors which recur at specific lags ➔ seasonal lag of target.
- ACF can demonstrate long decaying lags ➔ non-stationary time-series (i.e., because of trend) and will have the largest autocorrelations with short lags.
- Can de-trend and de-seasonalise the time series and then use the ACF to more clearly identify other useful lags.

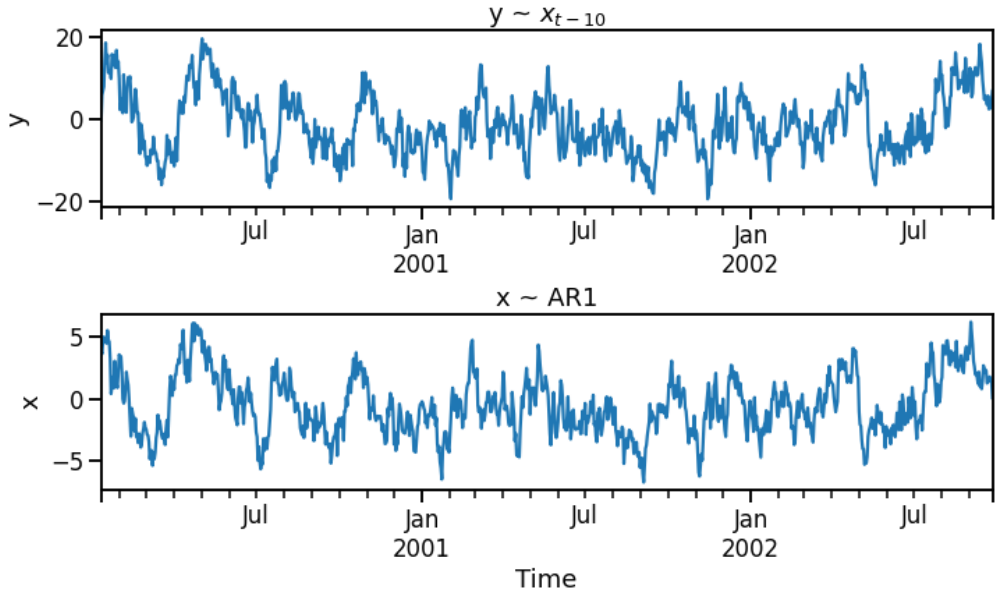
Partial autocorrelation function (PACF)

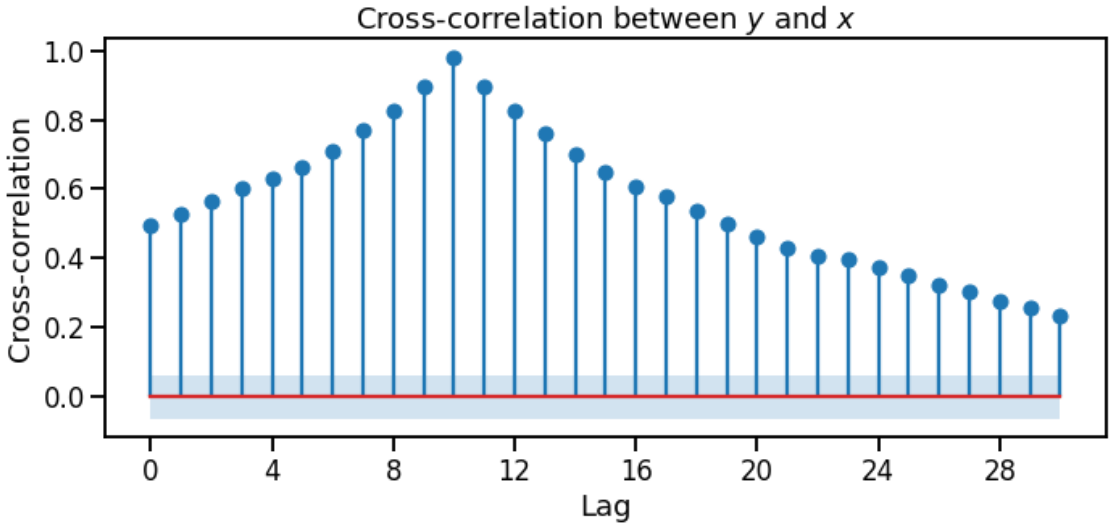




- PACF can better identify which shorter term lags to use because it accounts for correlation from previous lags.
- PACF can identify periodic behaviors which recur at specific lags ➔ seasonal lag of target.
- During the estimation of the PACF it is assumed the time series is stationary.
- Looking at both the ACF and PACF together can help determine whether to include a lag.

Cross-correlation function (CCF)

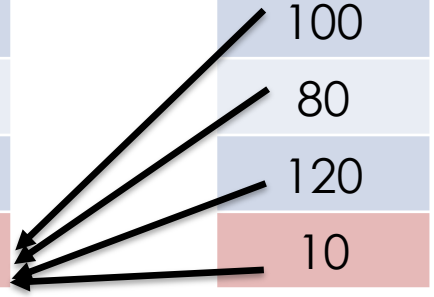




- CCF can identify which lags to use of another feature x_t to help predict the target y_t.
- During the estimation of the CCF it is beneficial to make the time series stationary.
- CCF only measures the linear correlation and will not identify non-linear relationships.

Distributed lags

Date	Sales	Ad spend
2020-02-12	23	100
2020-02-13	30	80
2020-02-14	35	120
2020-02-15	?	10



Consider a variable that has an impact distributed over time (e.g., advertising spend).

Distributed lags

Date	Sales
2020-02-12	23
2020-02-13	30
2020-02-14	35
2020-02-15	?

Ad spend	Ad spend Lag 1	Ad spend Lag 2	Ad spend Lag 3
100	NaN	NaN	NaN
80	100	NaN	NaN
120	80	100	NaN
10	120	80	100

Distributed lags are multiple lags of a variable that has an impact distributed over time.

Summary

Lags of the target and other features can create predictive features for forecasting.

Recent lags, seasonal lags, and distributed lags are different types of lag which can be helpful for forecasting.

Domain knowledge, feature selection & modelling, and time-series correlation methods can help decide which lags to use.