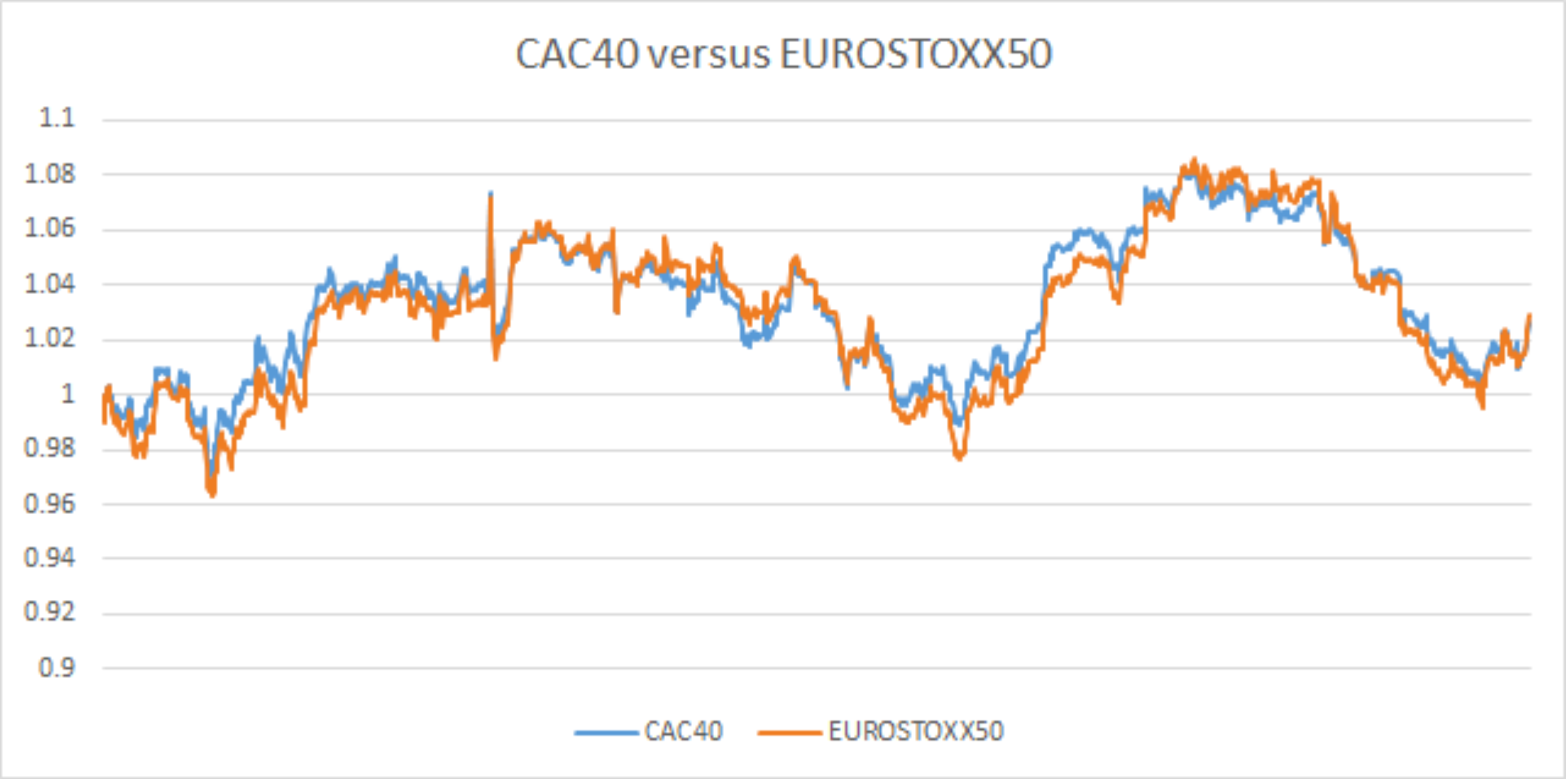
Algorithmic Trading

Dr. Javier Sabio González
Head of Advanced Analytics & Algorithmic Trading, BBVA C&IB

Topics

1. Markets Macrostructure
2. Markets Microstructure
3. Algorithmic Trading Fundamentals
4. Algorithmic Execution
5. Algorithmic Market-Making
6. **Algorithmic Investment**
7. The Future of Algorithmic Trading

6. Algorithmic Investment



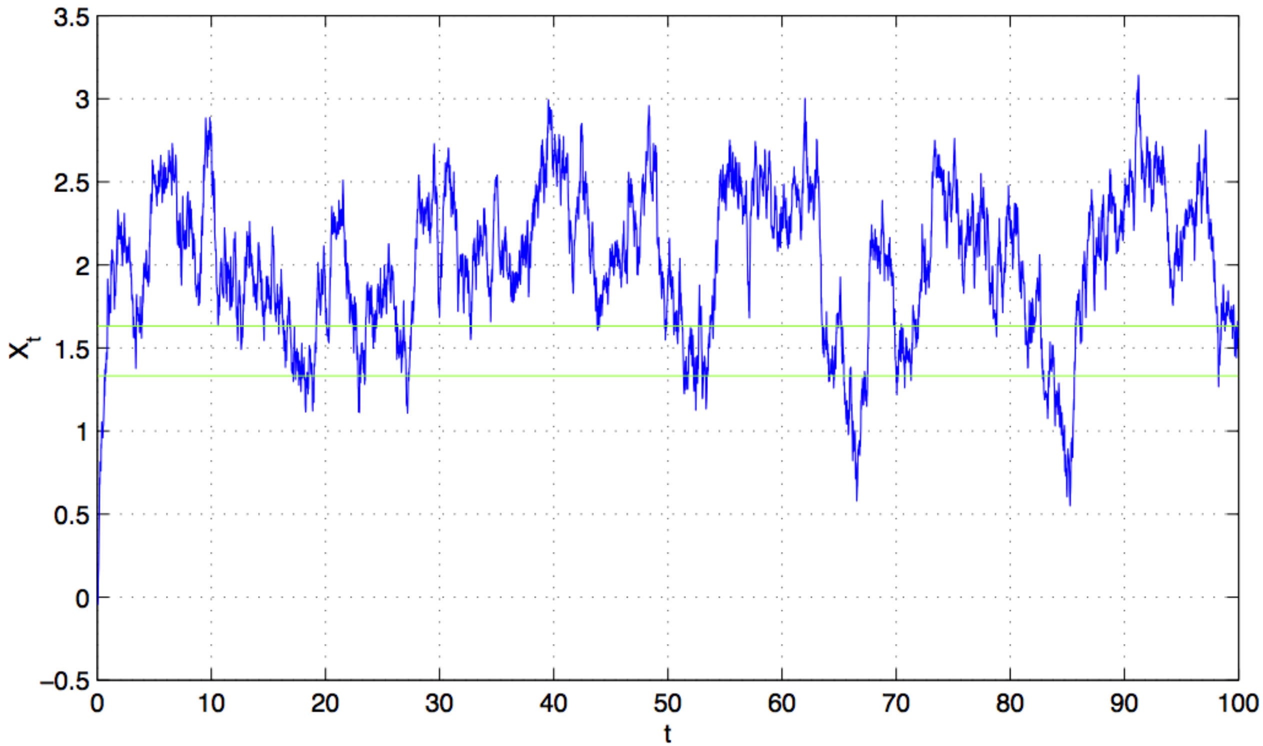
1. Mean-reversion strategies

Aim: strategies that try to detect deviations in market conditions that will likely revert to average conditions.

Remarks:

- Mean reversion can be analysed for single instruments or portfolios of them
- A very typical strategy is that of "pairs trading", that tries to detect mean-reversion in a portfolio composed of two correlated instruments
- It requires a statistical understanding of the concept of mean reversion and techniques to test for it

What is mean-reversion?



- Intuitively: the price of the instrument tends to revert to its long-term average
- Mathematically: the change of the price series in the next period is proportional to the difference between the mean price and the current price

What is mean-reversion?

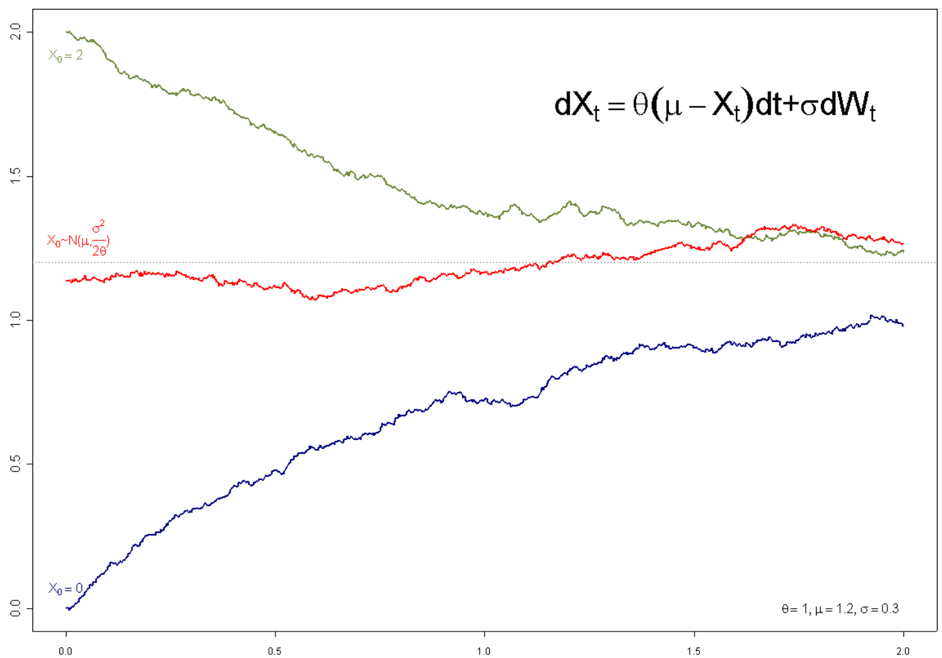
- Orstein-Uhlenbeck processes: simple stochastic model that shows mean-reversion

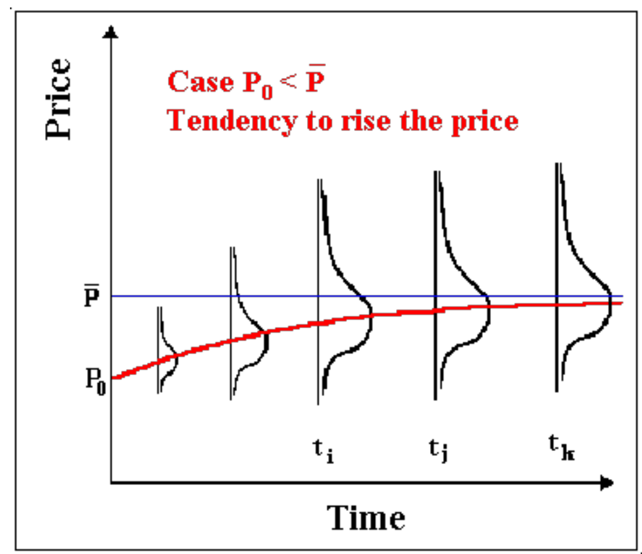
$$dx_t = \theta(\mu - x_t)\,dt + \sigma\,dW_t$$

$$\mathrm{E}(x_t) = x_0 e^{-\theta t} + \mu(1 - e^{-\theta t})$$

$$\mathrm{cov}(x_s, x_t) = \frac{\sigma^2}{2\theta}\left(e^{-\theta|t-s|} - e^{-\theta(t+s)}\right)$$

Ornstein - Uhlenbeck





Source:Wikipedia

What is mean-reversion?

- Discrete version:

$$\Delta x_k = \theta(\mu - x_k)\Delta t + \sigma \epsilon_k \Delta t$$

- This is a particular case of a AR(1) time-series model:

$$x_{k+1} = c + a x_k + b \varepsilon_k$$

$$c = \theta \mu \Delta t$$
$$a = 1 - \theta \Delta t$$
$$b = \sigma \sqrt{\Delta t}$$

Detecting mean-reversion: Dickey-Fuller Test

- Tests if the current price level has explanatory power over the price change
- Idea: fit an AR(1) time-series model to the empirical price time-series and test the hypothesis that the linear term is equal to one (zero for the model in finite differences)
- It uses the same statistical test as for the significance of the linear ("beta") term in a regression of the form:

$$\Delta x_k = \alpha + \beta x_k$$

- Test statistic:

$$DF = \frac{\hat{\beta}}{SE(\hat{\beta})}$$

- Null hypothesis: $\beta = 0$

If null hypothesis is rejected then the series shows mean-reversion with statistical significance

Critical values for Dickey–Fuller *t*-distribution.

	Without trend		With trend	
Sample size	1%	5%	1%	5%
T = 25	–3.75	–3.00	–4.38	–3.60
T = 50	–3.58	–2.93	–4.15	–3.50
T = 100	–3.51	–2.89	–4.04	–3.45
T = 250	–3.46	–2.88	–3.99	–3.43
T = 500	–3.44	–2.87	–3.98	–3.42
T = ∞	–3.43	–2.86	–3.96	–3.41

Detecting mean-reversion: Augmented Dickey-Fuller Test

- Generalises Dickey-Fuller to a more complicated time series regression:

$$\Delta x_k = \alpha + \beta x_k + \gamma t + \delta_1 \Delta x_{k-1} + ... + \delta_n \Delta x_{k-n}$$

- Test statistic: $DF = \dfrac{\hat{\beta}}{SE(\hat{\beta})}$

- Null hypothesis: $\beta = 0$

If null hypothesis is rejected then the series shows mean-reversion with statistical significance

Critical values for Dickey–Fuller *t*-distribution.				
	Without trend		With trend	
Sample size	1%	5%	1%	5%
T = 25	–3.75	–3.00	–4.38	–3.60
T = 50	–3.58	–2.93	–4.15	–3.50
T = 100	–3.51	–2.89	–4.04	–3.45
T = 250	–3.46	–2.88	–3.99	–3.43
T = 500	–3.44	–2.87	–3.98	–3.42
T = ∞	–3.43	–2.86	–3.96	–3.41

Detecting mean-reversion: Hurst Exponent

- Statistic that quantifies the long-term memory of a time series, i.e., the speed at which temporal autocorrelations decay
- If z = log (x), then we find the exponent that best adjusts the structure function:

$$\langle |z(t+\tau) - z(t)|^2 \rangle \sim \tau^{2H}$$

where τ is an arbitrary positive time lag.

- Cases:
  - if H = 0.5, then the price series diffuses as a geometrical random walk
  - If H < 0.5, the price series diffuses slower than a random walk, a signature of a mean - reverting behaviour
  - If H > 0.5, the price series diffuses faster than a random walk, a signature of a trending behaviour

Half-life of a mean-reverting process

- Recall the Orstein-Uhlenbeck process:

$$dx_t = \theta(\mu - x_t)\,dt + \sigma\,dW_t$$

- If the time-series can be fitted with such a model, then its expected value is:

$$\mathrm{E}(x_t) = x_0 e^{-\theta t} + \mu(1 - e^{-\theta t})$$

which decays exponentially to the long-term mean

- The half-life is the time necessary for the process to decay to 50% of the initial distance to the long-term mean

$$E[x_{t_{1/2}}] = \mu + \frac{(x_0 - \mu)}{2} = x_0 e^{-\theta t_{1/2}} + \mu(1 - e^{-\theta t_{1/2}})$$

$$t_{1/2} = \frac{\log 2}{\theta}$$

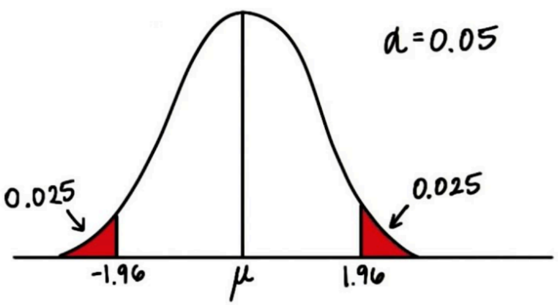
- The half-life provides a natural time-scale for the mean-reversion to occur, and therefore is a very useful magnitude to consider when designing a trading strategy

A mean-reversion strategy

- Identify a financial instrument with mean-reversion using ADF and/or Hurst
- Calculate the half-life of mean reversion
- Over a half-life time window, use the price series of the instrument to estimate the standard deviation and the mean. Depending on the half-life, different time scales for the returns have to be considered: daily, hourly, minutely…
- Monitor the z-score of the current value of the price of the instrument:

$$z_t = \frac{x_t - \mu_x}{\sigma_x}$$

- Entry and exit rules:



Entry (long) Exit Entry (short)

- Backtesting: historical p&l histograms, cumulative returns, and (daily, …) Sharpe ratio

$$\text{Sharpe ratio} = \frac{r_s - r_f}{\sigma_s}$$

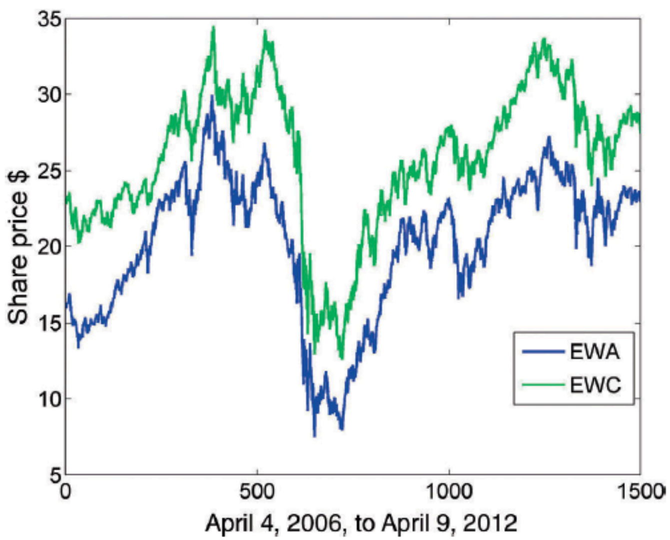
TABLE 1.1 Critical Values for $\sqrt{n} \times$ Daily Sharpe Ratio

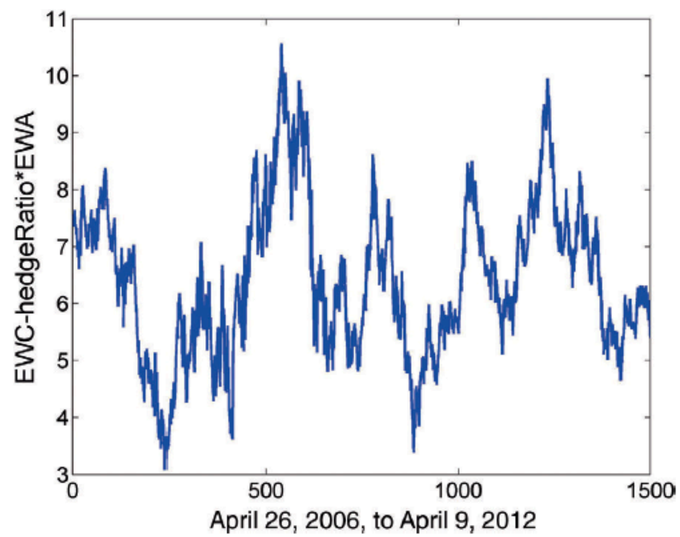
p-value	Critical values
0.10	1.282
0.05	1.645
0.01	2.326
0.001	3.091

Source: Berntson (2002).

Cointegration

- A set of time series is said to be cointegrated (or stationary) if there is a linear combination of them such as the resulting time series is mean-reverting
- The most common case is that of two time-series. When used in the context of trading strategies, this technique is the basis of the “pairs trading” strategy

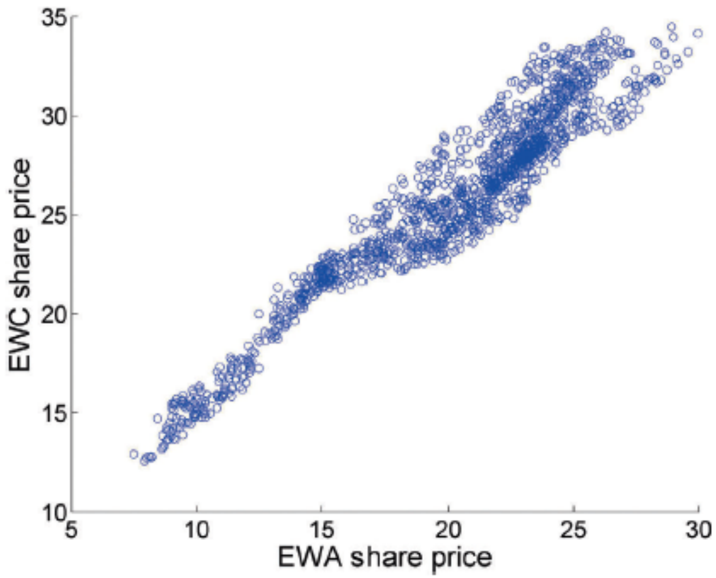




- Example (from Chan’s book): a linear regression between EWA and EWC is used to construct a mean-reverting portfolio. Both price series are cointegrated

Cointegrated Augmented Dickey-Fuller Test

- Procedure to find a cointegrated linear combination of two time series (if any), and test for the statistical significance of the mean-reversion:
  - Do a linear regression between time series one and time series two. Depending on which one is used as dependent / independent variable, two different and non-equivalent portfolios are obtained
  - Use the “beta” of the regression as the weight to combine the two series
  - Use the Augmented Dickey-Fuller test to check if mean-reversion of this portfolio is statistically significative



Johansen test

- Procedure to find a linear combination of several (two and more) time - series, if any, and test for cointegration.
- Recall that the Augmented Dickey Fuller test only works for two time-series
- The idea is similar to that of ADF, but for a vector where each component is a time series. The Johansen test fits the vector time-series to a Vector Autoregressive Model (VAR) and checks for the statistical significance of the minus one lag term (matrix A):

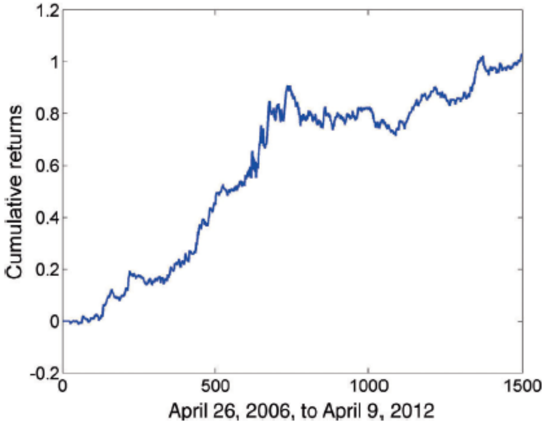
$$\Delta \mathbf{x_t} = \mu + A\mathbf{x_{t-1}} + \Gamma_1 \Delta \mathbf{x_{t-1}} + \ldots + \Gamma_p \Delta \mathbf{x_{t-p}} + \mathbf{w_t}$$

  - Idea: perform an eigenvalue decomposition of matrix A and tests for the rank *r* of the matrix. Null hypothesis is *r = 0*, meaning that there is no cointegration at all
  - If *r > 0*, one can use the eigevenctors of the matrix to form cointegrating portfolios, the one with the largest eigenvalue producing the best candidate
  - An advantage of the Johansen test with respect to cointegrated ADF, even for two - time series, is that the linear combination coefficients are calculated as part of the test, which can result in less false negatives of the null hypothesis.

Mean-reversion strategies for portfolios & Pairs Trading

- Use the Johansen test to find if there is a linear combination of a set of N instruments that is cointegrated
- Apply the mean-reversion strategy over the resulting portfolio
  - Calculate the half-life of mean-reversion
  - Calculate the rolling mean and standard deviation of the portfolio time series over the half-life time scale
  - Monitor the z-score
  - Use as entry signal when the z-score is beyond a certain percentile. Normally two sigmas are used ($z > 1.96$ or $z > -1.96$)
  - Use as exit signal when z gets close to zero.
  - Backtesting: historical, calculate p&l histogram, cumulative returns, Sharpe ratio

  - Example (from Chen's book): mean-reverting strategy over portfolio of EWA-EWC-IGE



Kalman Filter for cointegrated portfolios

- Idea: portfolio weights calculated with the Johansen test or cointegrated ADF are not stationary themselves, and vary over time. This means the latest points in the time-series should be weighted more
- Some possibilities
  - Rolling window of data: needs to find optimal window, drops completely far data
  - Moving average: needs to find optimal decay, arbitrary weighting
- A better alternative is a Kalman filter, which models explicitly the time-dependence of the portfolio weights. These are treated as latent variables (non-observable) and the new observations on the time series are used to update optimally the value of the latent variables to be consistent with the new information

Kalman Filter for Pairs Trading

- We elaborate the idea for the case of two instruments. In the most simple case, we model the relationship between the two price series as a linear regression with zero intercept:
$$y_t = \beta_t x_t + \epsilon_t,\ \epsilon_t \sim N(0, \sigma_\epsilon^2)$$
- In this case, the portfolio weight beta is time-dependent, and the latent (not observable) variable that has to be dynamically estimated. What are observed are the prices of the instruments, x(t) and y(t).
- In the Kalman Filter framework, this is called the “measurement or observation equation”, since it relates the latent variable with the observation although with some white noise parametrised by epsilon.
- In order to capture the time dependency of the portfolio weight, a simple random walk model is used, giving rise to the following “state transition” equation:
$$\beta_t = \beta_{t-1} + \omega_t,\ \omega_t \sim N(0, \sigma_\omega^2)$$
- The Kalman Filter gives an optimal estimation of the portfolio weight that is updated with new observations. For that we need to estimate $\sigma_\epsilon^2, \sigma_\omega^2$ with historical data

Kalman Filter for Pairs Trading

- Thanks to the simple gaussian models used for the observation and transition equations, the optimal estimate of the portfolio weight conditional to an observation of the two prices has a closed form:

$$\beta_t = \beta_{t-1} + \frac{x_t \sigma_\omega^2}{x_t^2 \sigma_\omega^2 + \sigma_\epsilon^2}(y_t - \beta_{t-1} x_t)$$

- The Kalman Filter can be easily coded, although it is included in many standard packages due to its wide applicability.
- A pairs trading strategy using the Kalman Filter only differs from the cointegrated mean reversion strategy with pairs of instruments, in that portfolio weights are dynamically updated.
- In general empirical studies tend to show that using the Kalman Filter to estimate the portfolio weights improves the performance of the pairs trading strategy
- A more general model allows for a non-zero spread between time series:

$$y_t = \alpha_t + \beta_t x_t + \epsilon_t = \begin{pmatrix} \alpha_t & \beta_t \end{pmatrix} \begin{pmatrix} 1 \\ x_t \end{pmatrix} + \epsilon_t$$

$$\begin{pmatrix} \alpha_t \\ \beta_t \end{pmatrix} = \begin{pmatrix} \alpha_{t-1} \\ \beta_{t-1} \end{pmatrix} + \begin{pmatrix} \omega_t^\alpha \\ \omega_t^\beta \end{pmatrix}$$

2. Momentum strategies

Aim: strategies that try to detect and benefit from trends in the market, e.g. in prices

- We concentrate here on strategies that exploit short term (intraday) momentum
- Some typical sources of short-term momentum are:
  - Triggering of stop losses
  - New information: corporate actions, macroeconomic events…
  - Rebalancing of large funds
  - Index composition changes, for stocks entering or exiting the index
  - Flow imbalance in the order book
  - Artificial flow imbalance created by high-frequency trading strategies
- Typical intraday momentum strategies:
  - Opening gap strategy
  - News-driven momentum strategy
  - High-frequency momentum strategies: ratio trade, ticking, momentum ignition, stop hunting

Detecting momentum

- There are two types of momentum, similarly to mean-reversion:
  - Time-series momentum: past returns of a price series are positively correlated with future returns
  - Cross-sectional momentum: a linear combination of price series shows time-series momentum
- In order to detect momentum, there are several possibilities:
  - calculate the linear correlation between future and past returns. In this case a lag needs to be chosen. Use a hypothesis test with null hypothesis the absence of correlation
  - calculate the correlation between signs of future and past returns
  - calculate the hurst exponent. As we saw previously, $H > 0.5$ is and indication of momentum.
  - use the ADF test. If the null hypothesis is rejected and the linear coefficient is positive (instead of negative as per mean reversion), this is momentum signal

Opening gap strategy

- The gap in a price time series is the difference between the closing price and the opening price in consecutive days or over weekends
- The rationale for momentum seeking is the existence of stop loss orders that might be triggered in cascade after the opening with a large enough gap
- There are many gap strategies based on different heuristics. For a successful strategy, a strong momentum signal from the gap is necessary, for instance:
  - Not only look at closing versus opening prices, but compare with OHLC from previous days to measure the significance of the gap. Alternatively, one can use a histogram of historical gaps to assess it
  - Look for abnormal and large volumes at the opening. Again, use historical data to assess this. Other external signals like breaking news, earnings… can be helpful.
  - Once the momentum signal is positive, place entry and exit orders:
    - Levels calibrated using historical data
    - Use high and low in OHLC from some period after the

News-driven momentum strategy

- Exploits the potential slow diffusion of news into the prices of financial instruments, which can take seconds, minutes, hours or even days
- The typical news exploited in strategies are: central bank conferences (ECB, FED), earnings announcements, changes in ratings or stock recommendations, housing and unemployment announcements (USA, China, Germany,…), consumer confidence, …
- News driven strategies typically use sentiment indicators:
  - Sentiment indicators translate news into a numerical score or tag that classifies the relevance of the news and its positive or negative impact
  - They are built using natural language processing techniques. Open source libraries can be found to build sentiment indicators
  - In addition, many third parties like Bloomberg or RavenPack sell proprietary sentiment indicators that can be incorporated into trading strategies
  - A strategy calibrates with historical data a threshold for the sentiment score that triggers momentum, and uses the real-time score as an entry signal.
- Other kind of strategies use “alternative data” to predict the outcome of the news and anticipate it: eg. looking at new insurance data to anticipate car sales

High-frequency momentum strategies

- Typically based on Limit Order Book data at high frequencies to predict short-term trends. For instance,:
  - the bid-ask volume imbalance is statistically correlated with a price movement in the opposite direction, e.g. if volume(bid) >> volume(ask): up trend
  - the order flow imbalance in a recent period (seconds, minutes) is also statistically correlated with short-term trends
- Many strategies try to exploit these statistical relationships. In order to profit on average, a large number of trials is necessary.
- **Ratio trade**: used in pro-rata based markets, when a short-term trend is expected, e.g., uptrend (downtrend), join immediately the bid (ask). A part of the order will be filled (pro-rata) and once the price move upwards (downwards), place order at best ask (bid) to get a profit.
- **Ticking or quote matching:** used in markets with bid/ask spread larger than one tick. When a short-term trend is expected, e.g. uptrend (downtrend), place the order at best bid (ask) plus (minus) one tick. Once is filled, sell at the best ask (bid) minus (plus) one tick.

High-frequency momentum strategies

- **Momentum ignition:** create the illusion of imbalance by placing a large order at the best bid (ask) and a small one a the best ask (bid). If other algorithms are monitoring imbalance they might fill our small order in the expectation of short-term trade. At that moment the large order is cancelled and replaced by a small one of the size of the one filled. When restoring balance, the algorithms that expected an up trend might now sell back at a loss, from which the strategy profits.
  - This strategy has many risks, both at the level of implementation (e.g. the possibility that our large order is actually filled, requiring us to unwind it probably at a lost), and from a regulatory point of view (market abuse)
- **Stop hunting:** look at support / resistance levels, where typically a lot of stop orders are placed. The strategy places relatively large buy (at resistance) or sell (at support) orders trying to trigger the stop orders, creating a cascade momentum effect. Once the momentum starts to fade down, the strategy unwinds the position at a profit.

3. Arbitrage strategies

Aim: strategies that try to detect misprices of financial instruments, for instance between derivatives and their underlyings, between prices of the same instrument in different markets, etc

Types:

- "True arbitrages": strategies that produce returns in excess of the risk-free return with zero risk
- "Statistical arbitrages": relative value strategies with a positive expected return and an acceptably small potential loss

True-arbitrages

- Between derivatives and underlyings, or between different derivatives, as far as the arbitrage is relatively model free:
  - Examples: Futures on an index vs basket of underlyings, bonds vs credit default swaps + interest rate swaps, treasury vs swap, future vs forward, put-call parity…
- Between instruments traded in different markets:
  - Examples: Telefonica in BME vs Chi-X, MX derivatives in MexDer Vs CME, …
- Between three pairs of currencies
  - Example: EURUSD, USDJYP, JYPEUR
- Notice: in most of the cases, there is a permanent basis between these instruments that cannot necessarily be exploited as a true-arbitrage. A very good understanding of markets and instruments is needed to prevent loses. Examples:
  - Liquidity basis: the different instruments/markets have different liquidity
  - Contractual basis: details on the mechanics of the instrument prevent a perfect arbitrage, as in credit default swaps vs bonds
  - Funding basis: cost of funding a short position, margin in a listed derivative…
- The way to cope with this basis is to analyse the time series of the arbitrage portfolios with mean reversion techniques: they become "statistical arbitrages"

Statistical-arbitrages (“StatArbs”)

- Mean reversion portfolios discussed previously are one kind of statistical arbitrages
- StatArbs however include a larger family of strategies, based on different models and domain expertise:
    - Risk-neutral model arbitrages: volatility strategies (options with different strikes and/or maturities, variance swaps, etc), term structure strategies (yield curve, swap curve, …), mortgage derivatives strategies
    - Machine-learning model arbitrages: popular recently, but very prone to overfitting and “garbage in, garbage out”
    - Capital structure arbitrage: taking long and short positions between different kind of instruments issued by a company: credit arbitrage (long/short CDS vs short/long stocks), convertible arbitrage (long convertible bonds vs short stocks)
- As we have seen, in practice “true-arbitrages” become statistical arbitrages, since liquidity, contractual issues, technology etc prevent the perfect arbitrage, and statistical techniques are required to monitor and signal the arbitrage opportunity