# Are Cryptos becoming alternative Assets?

Daniel Traian Pele, Niels Wesselhöfft, Wolfgang K. Härdle, Yannis Yatracos, Michalis Kolossiatis

International Research Training Group 1792
Ladislaus von Bortkiewicz Chair of Statistics
Humboldt–Universität zu Berlin

Department of Statistics and Econometrics
Bucharest University of Economic Studies

Department of Mathematics and Statistics
University of Cyprus, Nicosia

Yau Mathematical Sciences Center, Tsinghua University, Beijing, China

# Genus differentia approach

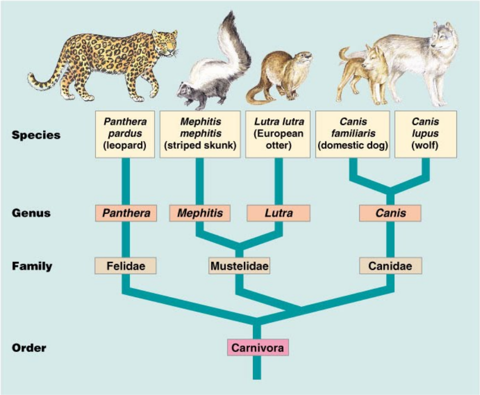





Figure: Genus differentia approach in biology

# Genus differentia approach

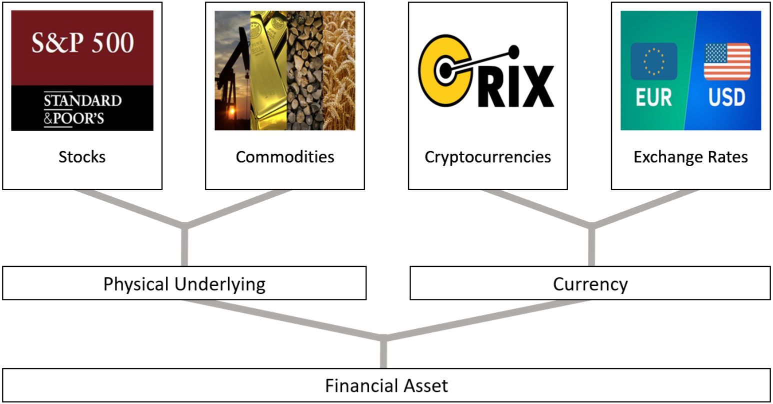


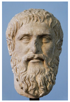
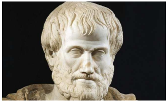

Figure: Genus differentia approach in finance

# Aim of classification

- Genotypic differentiation
  - Biology - the change in DNA sequences.
  - Finance - the underlying process of price manifestation.
- Phenotypic differentiation
  - Biology - classification based on behavior and features of a species.
  - Finance - classification based on statistical features of the price series.

# Motivation

- Question: What defines Cryptos?

- Plato: man is an upright, featherless biped, with broad, fat nails.
- Aristotle: definition of a species consists of genus proximum and differentia specifica.
- Goal: Define Cryptos in terms of their genus proximum and differentia specifica.
- Method: Find latent variables, to form groups of shared characteristics.
- Finding: Synchronic evolution, i.e. asymptotic speciation.
- Implication: Cryptos are a different species in the ecosystem of financial instruments.

# Outline

1. Motivation
2. Data and descriptives
3. Factor model
4. Explanation
5. Expanding window
6. Conclusion

# Literature review

- Dyhrberg (2016): BTC has similarities to both GOLD and the USD, being in between a currency and a commodity.
- Baur et al. (2018): BTC volatility and correlation characteristics are distinctively different compared to GOLD and USD.
- Härdle et al. (2018): BTC, XRP, LTC, ETH returns exhibit higher volatility, skewness and kurtosis compared to GOLD and S&P500 daily returns.
- Zhang et al. (2018): Cryptos presents heavier tails and higher Hurst exponent than the classical assets.
- Liu et al. (2019) developed a three-factor model using the CAPM approach and showed that the cross-sectional expected crypto returns can be captured by three factors: the market factor, the size factor and momentum factor.

# Data

- Sample: $n = 906$ assets.
- New asset class
  - Cryptos: $n_1 = 234$
- Classical assets
  - Stocks (S&P 500, EUROSTOXX 50, FTSE100): $n_2 = 635$
  - Exchange rates: $n_3 = 13$ List
  - Commodities (Bloomberg Commodity Index): $n_4 = 17$ List
  - Bonds: $n_5 = 6$
  - Real Estate: $n_6 = 2$
- Daily data from 03/01/2014 - 31/11/2020 (1740 trading days).

# Statistical assessment

- Return $X$ is a r.v. with cdf $F()$ from which $p = 24$ statistics are estimated.
- Moments of order $k \in \mathbb{R}^+$, $\mu_k = \mathsf{E}\{(X-\mu)^k\}$.
  - variance, skewness, kurtosis
- Tails: $\alpha \in \{0.005, 0.01, 0.025, 0.05, 0.95, 0.975, 0.99, 0.995\}$.
  - $Q_\alpha = \inf\{x \in \mathbb{R} : \alpha \leq F(x)\}$;
  - $CTE_\alpha = \begin{cases} \mathsf{E}\{X \mid X < Q_\alpha\}, & \alpha < 0.5 \\ \mathsf{E}\{X \mid X > Q_\alpha\}, & \alpha > 0.5 \end{cases}$
- Scaling and memory parameters
  - Alpha-stability ▸ Alpha-stability
  - autocorrelation coefficient $\rho_{it}(1)$
  - Hurst exponent $H_{it}$
  - $FIGARCH(1, d, 1)$ $d$ parameter

# Assets profile

| Indicator | Cryptos | Stocks | Bonds | Exchange rates | Commodities | Real Estate |
|---|---|---|---|---|---|---|
| $\sigma^2 \cdot 10^3$ | 26.442 | 0.400 | 0.745 | 0.028 | 0.388 | 0.146 |
| *Skewness* | 0.686 | −0.560 | −0.829 | −0.937 | −0.495 | −1.142 |
| *Kurtosis* | 35.163 | 20.004 | 35.623 | 31.949 | 15.089 | 17.045 |
| $Stable_\alpha$ | 1.342 | 1.602 | 1.460 | 1.722 | 1.662 | 1.696 |
| $Stable_\gamma$ | 0.047 | 0.009 | 0.009 | 0.003 | 0.009 | 0.006 |
| $Q_{5\%}$ | −0.183 | −0.027 | −0.029 | −0.008 | −0.027 | −0.018 |
| $Q_{2.5\%}$ | −0.251 | −0.038 | −0.039 | −0.010 | −0.036 | −0.024 |
| $Q_{1\%}$ | −0.366 | −0.054 | −0.057 | −0.013 | −0.050 | −0.034 |
| $Q_{0.5\%}$ | −0.485 | −0.071 | −0.082 | −0.015 | −0.065 | −0.041 |
| $CTE_{5\%}$ | −0.308 | −0.046 | −0.050 | −0.011 | −0.042 | −0.029 |
| $CTE_{2.5\%}$ | −0.404 | −0.060 | −0.066 | −0.014 | −0.054 | −0.038 |
| $CTE_{1\%}$ | −0.564 | −0.085 | −0.097 | −0.017 | −0.074 | −0.052 |
| $CTE_{0.5\%}$ | −0.719 | −0.106 | −0.130 | −0.021 | −0.091 | −0.067 |
| $Q_{95\%}$ | 0.190 | 0.027 | 0.028 | 0.008 | 0.027 | 0.018 |
| $Q_{97.5\%}$ | 0.276 | 0.036 | 0.041 | 0.010 | 0.035 | 0.023 |
| $Q_{99\%}$ | 0.422 | 0.051 | 0.060 | 0.013 | 0.050 | 0.029 |
| $Q_{99.5\%}$ | 0.581 | 0.068 | 0.078 | 0.015 | 0.065 | 0.034 |
| $CTE_{95\%}$ | 0.346 | 0.043 | 0.051 | 0.011 | 0.042 | 0.026 |
| $CTE_{97.5\%}$ | 0.467 | 0.056 | 0.069 | 0.013 | 0.053 | 0.031 |
| $CTE_{99\%}$ | 0.662 | 0.078 | 0.099 | 0.017 | 0.071 | 0.040 |
| $CTE_{99.5\%}$ | 0.843 | 0.095 | 0.130 | 0.019 | 0.085 | 0.048 |
| $\rho(1)$ | −0.116 | −0.043 | 0.150 | −0.001 | 0.024 | 0.070 |
| *Hurst* | 0.523 | 0.505 | 0.569 | 0.506 | 0.533 | 0.490 |
| *FIGARCHd* | 0.553 | 0.289 | 0.545 | 0.407 | 0.426 | 0.510 |

Table: Assets profile

# Factor analysis

- Estimate the correlation matrix for all variables.
- Factor extraction based on the correlation of the coefficients.
- Factor rotation.

# Correlation matrix

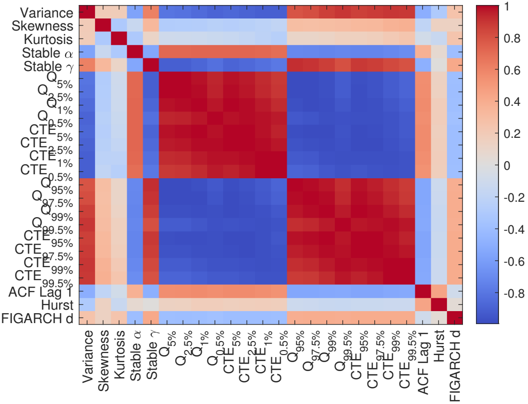



Figure: Correlation matrix of the statistical estimates. SFA_Cryptos

# Factor model

- Linear Factor model

$$X = QF + \mu + \varepsilon, \ \varepsilon \sim G() \tag{1}$$

- $X$ is the initial matrix of $p$ variables
- $Q$ is a matrix of the non-random loadings
- $F$ are the common $k$ factors ($k < p$)
- $\mu$ is the vector of the means of initial $p$ variables
- $\varepsilon$ is a matrix of the random specific factors
- Random vectors $F$ and $U$ are unobservable and uncorrelated

# Factors loadings and scree plot

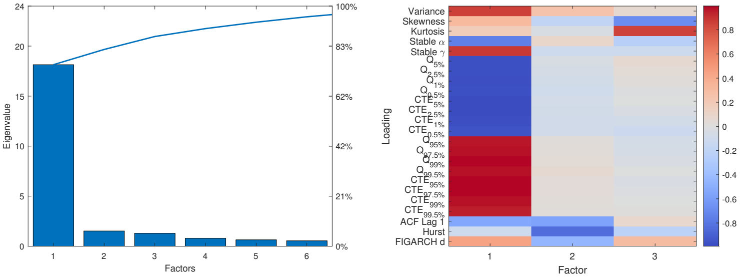



Figure: Scree plot and factors loadings. SFA_Cryptos

# Factor rotation

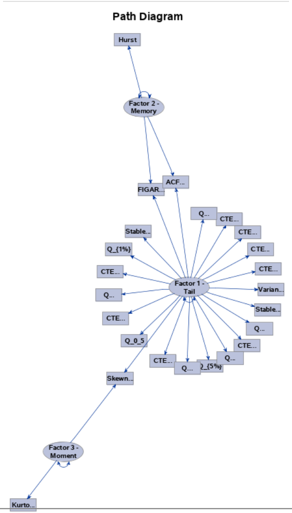



Figure: Path diagram. FA_Cryptos

# Mapping of the factors

1. Tail factor - 76% of the total variance
   - Alpha-stable parameters $S_\alpha$, $S_\gamma$
   - Lower and upper quantiles
   - Conditional tail expectations
   - Variance
2. Memory factor - 6% of the total variance
   - Hurst exponent
   - autocorrelation coefficient
   - $FIGARCH(1, d, 1)$ $d$ parameter
3. Moment factor - 6% of the total variance
   - Skewness
   - Kurtosis

# Factors

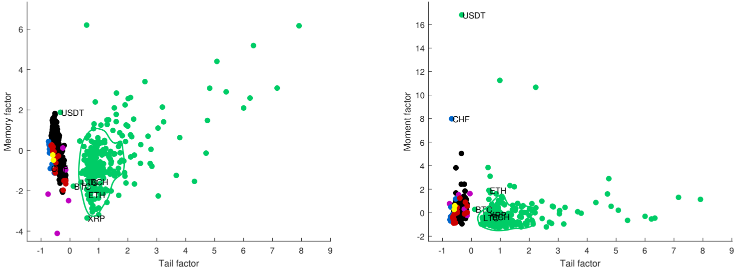



Figure: Assets projections on the factors space: (left) tail and memory factors; (right) tail and moment factors. SFA_Cryptos

# Factor explanation

- Classify between Cryptos and other asset classes
- Binary logistic regression for each factor $F_k$, $k \in \{1, 2, 3\}$

$$P(Y = 1) = \frac{\exp(\beta_0 + \beta_1 F_k)}{1 + \exp(\beta_0 + \beta_1 F_k)}, \tag{2}$$

$$Y = \begin{cases} 1, & \text{if crypto} \\ 0, & \text{if otherwise} \end{cases} \tag{3}$$

# Factor explanation

| Exogenous factor | Tail factor | Memory factor | Moment factor |
|---|---|---|---|
| Estimated $\beta_1$ | 15.450*** | -0.738*** | 0.116 |
| | (4.435) | (0.089) | (0.087) |
| $\widetilde{R}^2$ | 0.992 | 0.122 | 0.102 |

Note: Standard errors in (); ** denotes significance at 95% confidence level.

$$\widetilde{R}^2 = \frac{1 - \left\{ \frac{L(0)}{L(\widehat{\beta})} \right\}^{\frac{2}{n}}}{1 - \{L(0)\}^{\frac{2}{n}}} \tag{4}$$

- $L(0)$ is the likelihood of the intercept-only model
- $L(\widehat{\beta})$ is the likelihood of the full model

# Support Vector Machines

- Finding a projection that maximizes margin in a hyperplane of the original data.
- No parametric assumptions on the underlying probability distribution function.
- 99.56% accuracy.
- Missclassified: Bitcoin (BTC) and Tether (USDT).

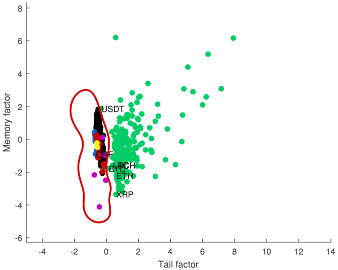



Figure: SVM ▸ SVM

# K-means clustering

- Projection of the clusters on the 3D space extracted trough Factor Analysis.
- 98.45% accuracy.

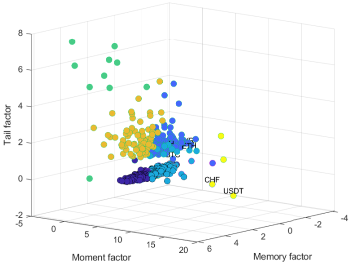



Figure: 3D. SFA_Cryptos

# Maximum Variance Components Split

- These method have goals to separate, respectively, the components of a structure like the types of assets herein, and clusters defined as the components of a mixture distribution.
- They are based on an unusual variance decomposition in between-group variations.

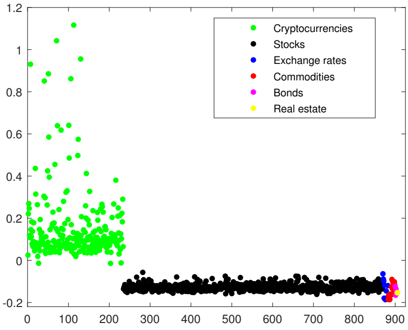
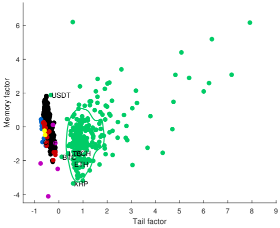



Figure: MVCS. VCS_cryptos

MVCS

# Video

- Expanding rolling window estimation
  - Starting window 03-01-2014 until 22-04-2016 (1/2 of the data)
  - Increases daily up to full window 31/11/2020
  - Kernel density contour level 0.05
- Clusters converge over time




DFA_cryptos

# Synchronic evolution

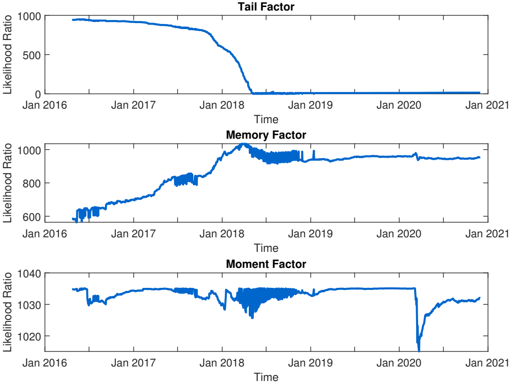



Figure: Likelihood Ratios for the binary logistic model, estimated for the period 03/01/2014- 31/11/2020. CONV_cryptos

# Impact of Covid-19

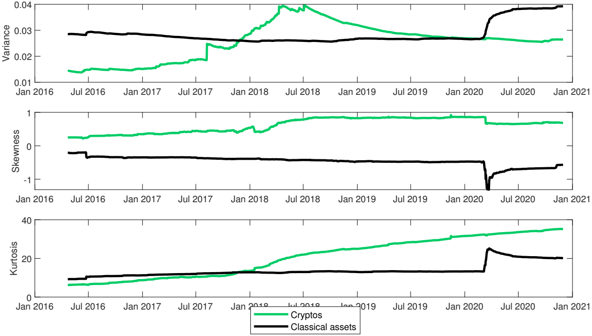



Figure: Variance, skewness and kurtosis dynamics by assets class. [1]

CONV_cryptos

[1] For classical assets, the variance is multiplied by $10^2$.

# Impact of Covid-19

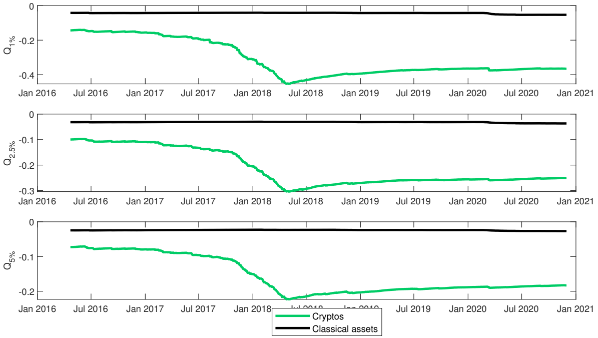



Figure: Quantile dynamics by assets class.

CONV_cryptos

# Market risk example

Table: : Average 1% VaR

| VaR method | Classical assets | Cryptos only | Mixed portfolio |
|---|---|---|---|
| Historical VaR | 3.01 | 16.34 | 4.37 |
| MVaR | 4.00 | 24.22 | 6.82 |
| Normal GARCH(1,1) | 2.04 | 12.24 | 3.21 |
| Student's t GARCH(1,1) | 2.22 | 14.39 | 3.56 |
| Normal GJR-GARCH(1,1,1) | 2.07 | 12.00 | 3.11 |
| Student's t GJR-GARCH(1,1,1) | 2.16 | 14.15 | 3.48 |

The average 1% VaR is reported in percentage terms; VaR was estimated using a rolling window of 250 trading days, for the period 03/01/2014-30/11/2020.

# Market risk example

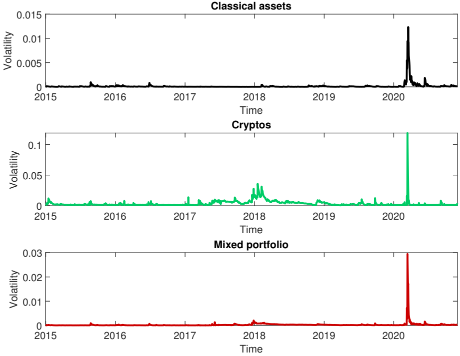



Figure: Estimated volatilities from Student's t GJR-GARCH(1,1,1) model, using a rolling window approach VaR_Cryptos

# Policy implications

- Given Cryptos' unpredictable and highly volatile behaviour, investors may be exposed to higher risks than investing in classical assets.
- Cryptos can be seen as an alternative for portfolio diversification, if investors are looking for higher compensation from riskier assets.
- Cryptos may not be suitable for risk-averse investors, especially in bear market circumstances.
- Conventional inference based on normal distribution appears to be inappropriate when it comes to the prudential treatment of Cryptos.
- Cryptos may require extra attention and monitoring, as their high volatility could jeopardize overall financial stability.

# Conclusion

- Financial perspective
  - Main statistical difference between Cryptos and other asset classes: tail behavior.
  - Moments and memory are of subliminal importance.
  - Nonlinear classification with SVM provides proficient results for risk analysts and regulators.
  - Cryptos are completely separated by the other types of assets, as proved by Maximum Variance Components Split method.
- Biological perspective
  - Speciation takes time to form distinct species, which potentially evolve further away from each other.
  - Cryptos establish themselves as unique asset classes.

# Are Cryptos becoming alternative Assets?

Daniel Traian Pele, Niels Wesselhöfft, Wolfgang K. Härdle, Yannis Yatracos, Michalis Kolossiatis

International Research Training Group 1792
Ladislaus von Bortkiewicz Chair of Statistics
Humboldt–Universität zu Berlin

Department of Statistics and Econometrics
Bucharest University of Economic Studies

Department of Mathematics and Statistics
University of Cyprus, Nicosia

Yau Mathematical Sciences Center, Tsinghua University, Beijing, China

# Exchange rates

▸ Data

1. EUR/USD Euro
2. JPY/USD Japanese Yen
3. GBP/USD Great Britain Pound
4. CAD/USD Canada Dollar
5. AUD/USD Australia Dollar
6. NZD/USD New Zealand Dollar
7. CHF/USD Swiss Franc
8. DKK/USD Danish Krone
9. NOK/USD Norwegian Krone
10. SEK/USD Swedish Krone
11. CNY/USD Chinese Yuan Renminbi
12. HKD/USD Hong Kong Dollar
13. INR/USD Indian Rupee

# Commodities

Data

1. WTI Crude oil USCRWTIC Index
2. Natural Gas NGUSHHUB Index
3. Brent oil EUCRBRDT Index
4. Unleaded Gasoline RBOB87PM Index
5. ULS Diesel DIEINULP Index
6. Live cattle SPGSLC Index
7. Lean hogs HOGSNATL Index
8. Wheat WEATTKHR Index
9. Corn CRNUSPOT Index
10. Soybeans SOYBCH1Y Index
11. Aluminum LMAHDY Comdty
12. Copper LMCADY Comdty
13. Zinc ZSDY Comdty
14. Nickel CKEL Comdty
15. Tin JMC1DLTS Index
16. Gold XAU Curncy
17. Silver XAG Curncy
18. Platinum XPT Curncy
19. Cotton COTNMAVG Index
20. Cocoa MLCXCCSP Index

# Lévy-Stable distributions

- Fourier transform of characteristic function $\varphi_X(u)$

$$S(X \mid \alpha, \beta, \gamma, \delta) = \frac{1}{2\pi} \int \varphi_X(u) \exp(-iuX) du$$

- Characteristic function representation, $0 < \alpha < 2, \alpha \neq 1$

$$\log \varphi_X(u) = iu\delta - \gamma |u|^{\alpha} \{1 + i\beta (u/|u|) \tan(\alpha\pi/2)\} \tag{5}$$

- Stability or invariance under addition

$$n \log \varphi_X(u) = iu(n\delta) - (n\gamma)|u|^{\alpha} \{1 + i\beta (u/|u|) \tan(\alpha\pi/2)\}$$

- Limiting distribution of $n$ i.i.d. stable r.v., $0 < \alpha \leq 2$
GCLT (Gnedenko and Kolmogorov, 1954)

$$n^{-\frac{1}{\alpha}} \sum_{i=1}^{n} (X_i - \delta) \xrightarrow{\mathcal{L}} S(\alpha, \beta, \gamma, 0) \tag{6}$$

# Linear Discriminant Analysis

- Let $X_i \sim N(\mu_i, \Sigma_i)$ belonging to class $\omega_i$, $\Sigma_i = \Sigma_j$
- Project samples $X$ onto a line $Y = w^\top X$
- Select the projection that maximized the separability
- Maximize normalized, squared distance in the means of the classes

$$w^* = \arg\max_w \frac{|w^\top(\mu_i - \mu_j)|^2}{s_i^2 + s_j^2}, \tag{7}$$

$$s_i^2 = \sum_{x_i \in \omega_i} (w^\top x_i - w^\top \mu_i)^2 = w^\top S_i w \tag{8}$$

- Linear Discriminant of Fisher (1936)

$$w^* = S_W^{-1}(\mu_i - \mu_j), \;\; S_W = S_i + S_j \tag{9}$$

▸ LDA

# Support Vector Machines

- Given training data set $D$ with $n$ samples and 2 dimensions

$$D = (X_1, Y_1), \ldots (X_n, Y_n),$$
$$X_i \in \mathbb{R}^2, \quad Y_i \in [0,1]$$

- Finding a hyperplane that maximizes the margin

$$\min_{w,b} \frac{1}{2} \|w\|^2$$
$$\text{s.t.} \quad Y_i \left( w^\top X_i + b \right) \geq 1,$$
$$i = 1, \ldots, n$$

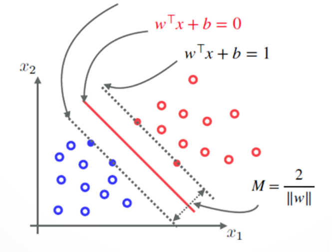



Figure: SVM

# Variance Component Split

- Consider the groups $X_{(1)}, \ldots, X_{(i)}$ and $X_{(i+1)}, \ldots, X_{(n)}$ with averages, respectively, $\overline{X}_{[1,i]}$ and $\overline{X}_{[i+1,n]}$, $i = 1, ..., n-1$, then

$$\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2 = \sum_{i=1}^{n-1}\frac{i(n-i)}{n^2}(\overline{X}_{[i+1,n]} - \overline{X}_{[1,i]})(X_{(i+1)} - X_{(i)}). \quad (10)$$

- The relative contribution of the groups $X_{(1)}, ..., X_{(i)}$ and $X_{(i+1)}, ..., X_{(n)}$ in the sample variability:

$$W_i = W_i(X_1, \ldots, X_n) = \frac{i(n-i)}{n}\frac{(\overline{X}_{[i+1,n]} - \overline{X}_{[1,i]})(X_{(i+1)} - X_{(i)})}{\sum_{i=1}^{n}(X_i - \overline{X})^2} \quad (11)$$

- Index $\mathcal{I}_n = \max\{W_i, i = 1, \ldots, n-1\}$ determines two potential clusters or parts of a structure and is based on averages and inter-point distances.

# Maximum Variance Component Split

- The Maximum Variance Component Split (MVCS) method compares known components of a structure, *e.g.* Cryptos herein, with data splits for a set of unit projection directions $\mathcal{D}_M$ usually determined by $M$ positive equidistant angles of $[0, \pi]$; *e.g.* when $r = 2$ and $M = 3$ the angles used are $\pi/3, 2\pi/3, \pi$.
- When one of the data split along projection direction a coincides with a component of the structure we have complete separation of this component along a.
- A set of projection directions $\mathcal{D}_M$ can be

$$(\Pi_{l=1}^{r} cos\theta_l,\ \ sin\theta_1 \Pi_{l=2}^{r} cos\theta_l, \dots,\ \ sin\theta_{r-1} cos\theta_r,\ \ sin\theta_r), \qquad (12)$$

where $\theta_l$ takes values in $\{\frac{m\pi}{M}, m = 1, \dots, M\}, l = 1, \dots, r$.

▸ MVCS