Independencia estadística: El Análisis técnico de mercados financieros como caso de apofenia social

Javier Arriero Pais

Agenda

1. Definición
2. Predicción de series temporales
3. Técnicas para medir la dependencia estadística
4. Análisis técnico
5. Conclusiones

Definición

- $P(X, Y) = P(X)P(Y)$
- $\varphi_{X,Y}(s, t) = \varphi_X(s)\varphi_Y(t)$
- $MI(X, Y) = 0$
- $corr(f(X), g(Y)) = 0, \forall f \in C_X,\ g \in C_Y$

Definición probabilística

Independencia

$$P(X, Y) = P(X)P(Y)$$

Condicionar no afecta

$$P(X|Y) = P(X)$$
$$P(Y|X) = P(Y)$$

Definición armónica

Independencia

$$\varphi_{X,Y}(s,t) = \varphi_X(s)\varphi_Y(t)$$

Transformada de Fourier de la función de densidad de probabilidad

$$\varphi_X(s) = \int_{\mathbf{X}} e^{-isx} p(x) dx$$

Definición bajo la teoría de la información

Independencia

$$MI(X,Y) = 0$$

La información que aporta una variable sobre la otra es cero

$$MI(X,Y) = \int_{\mathbf{Y}} \int_{\mathbf{X}} \log \frac{p(x,y)}{p(x)p(y)} p(x,y) dx dy$$

Definición funcional

Independencia

$$corr(f(X), g(Y)) = 0, \forall f \in C_X, g \in C_Y$$

No existe un par de transformaciones que hagan correlar las variables

$$corr(f(X), g(Y)) = \frac{Cov(f(X), g(Y))}{\sqrt{Var(f(X))}\sqrt{Var(g(Y))}}$$

Predicción de series temporales

Procesos estocásticos

$$...X_{t-2}, X_{t-1}, X_t, X_{t+1}, X_{t+2}, ...$$

Modelos autorregresivos

$$X(t+1) = f(X(t), \epsilon)$$

Información del futuro en el presente

Hipótesis

- $X(t)$ contiene información sobre $X(t+1)$
- $X(t)$ y $X(t+1)$ no son independientes

Rechazar uno de los siguientes

- $P(X(t+1), X(t)) = P(X(t+1))P(X(t))$
- $\varphi_{X(t+1),X(t)}(r,s) = \varphi_{X(t+1)}(r)\varphi_{X(t)}(s)$
- $MI(X(t+1), X(t)) = 0$
- $corr(f(X(t+1)), g(X(t))) = 0,\ \forall f \in C_{X(t+1)},\ g \in C_{X(t)}$

Clonado del repositorio

```
git clone https://github.com/analyticbastard/statistical-independence-
    financial-forecasting.git
```

Sesión práctica: Instalación de paquetes de CRAN y carga d

datos

Abrir el fichero `load-data.R`

Correlación lineal: Sesión práctica

Abrir el fichero `linear.R`

Kernel CCA

$$corr(f(X), g(Y)) = 0$$

Generar infinitas combinaciones de funciones en C_X y C_Y para probar

- Imposible

Usar el kernel trick

- Los RKHS son suficientemente grandes en C (densos en L^2)
- Pueden extenderse métodos basados en producto interior (PCA, CCA, SVM...)

Reproducing Kernel Hilbert Space

Espacio de funciones generado de manera especial

- $k(x, y)$ p.s.d. es el kernel del espacio $\mathbf{H}_k$. Se fija un parámetro del kernel: $k_x = k(x, \cdot)$. $f \in \mathbf{H}_k$, $f(x) = \langle f, k_x \rangle = \sum_{i=0}^{\infty} \alpha_i k(x_i, x)$: Delta de Dirac (funcional de evaluación puntual) $k_x \in \mathbf{H}_k$
- Para cada dato x, $\mathbf{H}_k$ asocia una imagen $\Phi : \mathbf{X} \to \mathbf{H}_k, \Phi(x) = k_x$.
- Aproximación finita: Para un conjunto de datos $(x_1, x_2, \ldots, x_n)$, K es la matriz kernel evaluando la función k en los pares de puntos, $K_{ij} = k(x_i, x_j)$

Correlaciones canónicas

Buscar subespacios ortogonales de máxima correlación

- Subespacio de máxima correlacion:

$$\max_{\alpha_1 \in \mathbf{R}^N, \beta_1 \in \mathbf{R}^M} \operatorname{corr}(X\alpha_1, Y\beta_1)$$

- Siguiente par: igual, sujeto a $\alpha_1 \perp \alpha_2$ y $\beta_1 \perp \beta_2$
- X e Y pueden tener diferente dimensión

Correlaciones canónicas con kernel trick

$$\max_{\alpha_1 \in \mathbf{R}^N, \beta_1 \in \mathbf{R}^M} \frac{\alpha^T K_x K_y \beta}{\sqrt{\alpha^T K_x^2 \alpha}\sqrt{\beta^T K_y^2 \beta}}$$

KCCA: Sesión práctica

Abrir el fichero `kcca.R`

ForeCA

ForeCA

Forecastable Components Analysis, dada una serie temporal, ForeCA calcula un componente predecible y una serie residuo de ruido ortogonal a este, para lo que usa la entropía de Shannon de una métrica particular (aproximación de la información mutua)

`https://arxiv.org/abs/1205.4591`

ForeCA: Sesión práctica

Abrir el fichero `foreca.R`

Forma general

$$g(X_{t+n_1}, X_{t+n_2}, \ldots, X_{t+n_N}) = f(X_{t-i_1}, X_{t-i_2}, \ldots, X_{t-i_L})$$

Pero...

... hemos visto que las X_j son (cuasi) independientes entre ellas

Conclusiones

- Ni el análisis técnico ni una red neuronal pueden predecir futuros precios
- No free lunch ilustrado estadísticamente
- Para un problema de ML: Pensar antes de disparar
- $...X_{t-2}, X_{t-1}, X_t, X_{t+1}, X_{t+2}, ...$ es estacionario
- ML para rebalanceo de portfolio: Más allá de Markowitz
- Jim Simmons, Rentec, Medallion Fund