Arbitrage
จากแนวคิดสู่การปฏิบัติ

Part V: Statistical Arbitrage (บท 18-21)

Pairs Trading, Mean Reversion, Factor Models, ML Overview

บทที่ 18: Pairs Trading

Stat Arb เริ่มต้นจากแนวคิดง่ายๆ: **"หุ้น 2 ตัวที่เคลื่อนไหวด้วยกัน → เมื่อแยกจากกัน → เข้าเทรด"**

18.1 Correlation vs Cointegration

	Correlation	Cointegration
วัดอะไร	ทิศทางเหมือนกันไหม	**ระยะห่าง**คงที่ไหม
ปัญหา	2 หุ้นอาจ drift ออกจากกันตลอดกาล	Spread กลับหาค่าเฉลี่ย
ตัวอย่าง	BTC กับ ETH: correlated แต่ ratio เปลี่ยนเรื่อยๆ	Coke vs Pepsi: spread mean-reverts
ใช้กับ Arb?	ไม่พอ!	**ต้องใช้ตัวนี้!**

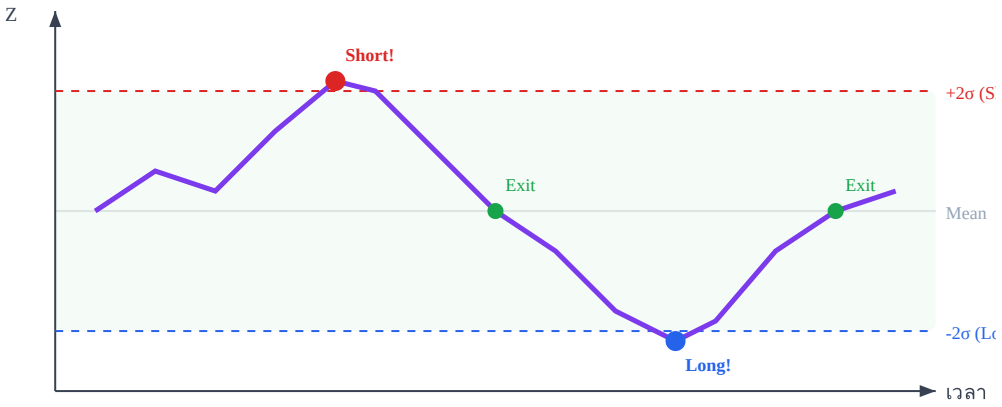
Engle-Granger Cointegration Test

1. Regress A vs B: $A = \alpha + \beta \times B + \varepsilon$
2. ทดสอบ residuals (ε) ว่า stationary ไหม (ADF test)
3. ถ้า ε stationary → A กับ B cointegrated → **Spread = A - β×B จะ mean-revert!**

18.2 Trading the Spread

```
Spread = Price_A - β × Price_B
Z-score = (Spread - Mean) / StdDev

Entry: Z > +2 → Short Spread (Short A, Long β×B)
Entry: Z < -2 → Long Spread (Long A, Short β×B)
Exit: Z ≈ 0 → Close ทั้งสองขา
Stop: |Z| > 4 → Cut loss (structural break?)
```


Z
Short!
+2σ (S
Exit
Exit
Mean
-2σ (Lo
Long!
เวลา

บทที่ 19: Mean Reversion

19.1 Ornstein-Uhlenbeck Process

เปรียบเทียบ: "ยางยืด"

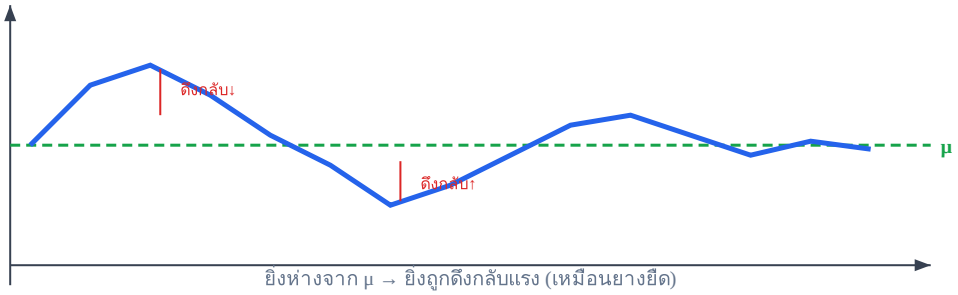
ลองนึกว่า spread ถูกผูกด้วยยางยืดกับค่าเฉลี่ย μ — ยิ่งดึงออกไกล ยิ่งถูกดึงกลับแรง

θ = ความแข็งของยาง (สูง = กลับเร็ว), σ = ลมที่พัดให้แกว่ง (noise)

```
dX = θ(μ - X)dt + σdW

θ = speed of mean reversion (ยิ่งสูง ยิ่งกลับเร็ว)
μ = long-run mean (ค่าที่ดึงดูดกลับ)
σ = volatility

Half-life = ln(2) / θ ← "นานแค่ไหนกว่า spread จะกลับครึ่งทาง"
```



ตัวอย่าง

ถ้า Half-life = 5 วัน → spread กลับครึ่งทางใน 5 วัน → ดีมาก trade ได้

ถ้า Half-life = 60 วัน → ช้าเกินไป → tie up capital นาน → อาจไม่คุ้ม

19.2 Kalman Filter — Dynamic Hedge Ratio

ปัญหา: β (hedge ratio) เปลี่ยนตามเวลา → ใช้ rolling regression ไม่ดีพอ

Kalman Filter: อัพเดท β แบบ real-time ตามข้อมูลใหม่ → **adaptive pairs trading**

19.3 Regime Detection

เมื่อไร Mean Reversion ใช้ไม่ได้?

Trending market → spread ไม่ mean-revert → **ขาดทุนต่อเนื่อง**

Structural break → cointegration หาย (M&A, sector change)

ต้องมี **regime detection**: ถ้าตลาดเปลี่ยน → หยุดเทรด pairs

บทที่ 20: Factor-Based Statistical Arbitrage

20.1 จาก Pairs สู่ Portfolio

เทรด 1 คู่ → diversify ด้วยหลายคู่ → สร้าง **portfolio ของ pairs trades**
→ single-pair risk ลดลง → more consistent returns

20.2 Factor Models

```
Return_i = α_i + β₁×Market + β₂×Size + β₃×Value + ... + ε_i

α (alpha) = ส่วนที่ model อธิบายไม่ได้ = edge ที่เราหา!
β×Factor = systematic risk → hedge ออกได้
```

20.3 Market-Neutral Portfolio

```
Portfolio beta = 0 (ตลาดขึ้น/ลงไม่กระทบ)
Long stocks ที่มี positive alpha + Short stocks ที่มี negative alpha
→ กำไรจาก alpha เท่านั้น ไม่ขึ้นกับทิศตลาด [Lv.4 Structured]
```

PCA Factors (จากเล่มคณิตศาสตร์ บท 8)

ใช้ Eigenvalues ของ Covariance Matrix → หา principal components
PC1 ≈ market factor, PC2 ≈ sector rotation, PC3 ≈ size/value
→ Hedge ทีละ factor → เหลือแค่ alpha

บทที่ 21: Machine Learning ใน Stat Arb

21.1 Feature Engineering

สร้าง features จากข้อมูลดิบ: momentum, mean-reversion signals, vol ratios, order flow

21.2 Walk-Forward Cross-Validation

> **Time Series ≠ Regular Data!**
>
> ห้ามใช้ random K-fold → data leakage (ใช้อนาคตทำนายอดีต)
>
> ต้องใช้ **Walk-Forward**: train on past → test on next period → roll forward

21.3 Overfitting — ศัตรูตัวฉกาจ

```
ถ้า backtest Sharpe = 5.0 → 99% overfitted
ถ้า backtest Sharpe = 1.5 → อาจจริง → ทดสอบ out-of-sample

กฎ: Simple models > Complex models (ส่วนใหญ่)
Linear regression มักชนะ neural network ใน finance
```

เรื่องจริง: Pairs Trading — Royal Dutch Shell A vs B (1907-2005)

Shell มีหุ้น 2 classes (Dutch=A, English=B) สัดส่วน 60:40 → ราคาควรเป็น 60:40 ตลอด
จริงๆ: ratio เบี่ยงเบน ±15% เป็นประจำ → pairs traders ทำกำไรหลายทศวรรษ
แต่ **LTCM ก็เทรดคู่นี้แล้วเจ๊บ** เพราะ ratio กว้างขึ้นก่อน converge → margin call
บทเรียน: Pairs ที่ดีที่สุดก็มี timing risk — ต้อง size ให้รอดถึงวัน converge

ลองคิดดู (บท 18-21)

1. หุ้น A กับ B มี correlation = 0.95 → ใช้ pairs trading ได้ไหม? ทำไม?
2. Spread มี half-life = 3 วัน vs 90 วัน → อันไหนดีกว่าสำหรับ pairs trading?

เฉลย: 1. Correlation สูงไม่พอ! ต้องทดสอบ cointegration — ถ้าไม่ cointegrated spread อาจ drift ไปเรื่อยๆ 2. 3 วันดีกว่ามาก — กลับเร็วกว่า tie up capital น้อยกว่า

สรุป Part V (บท 18-21) — Part 6/9 ✓

✓ **Pairs Trading**: Cointegration > Correlation, Z-score entry/exit
✓ **OU Process**: θ = speed, Half-life กำหนดว่าคุ้มไหม
✓ **Kalman Filter**: Dynamic hedge ratio
✓ **Factor Models**: alpha = edge, beta = hedge ออกได้
✓ **Market-Neutral**: Long alpha + Short alpha, beta = 0
✓ **ML**: Walk-Forward CV, Overfitting = ศัตรู, Simple > Complex
✓ **ทั้งหมดเป็น Lv.4 [Heuristic]** — ไม่ใช่ arb จริง มี statistical edge

จบ Part V

Part VI: Merger & Event Arbitrage