Chapter. 03

시계열의 평활화

Kalman 필터

FASTCAMPUS
ONLINE
금융공학/퀀트

강사. 장순용

I 키포인트

- 시계열의 평활화.
- Kalman 필터.
- Kalman 필터 알고리즘.

Copyright FASTCAMPUS Corp. All Rights Reserved

I Kalman 필터: 개요

- 칼만 필터는 소위 “데이터 퓨젼 알고리즘” (data fusion algorithm)의 일종이다.
- 내재된 모형과 노이즈가 포함된 시계열을 융합하여 평활화 효과를 낸다.

Fast campus
Copyright FASTCAMPUS Corp. All Rights Reserved

I Kalman 필터: 개요

- 여기에서는 벡터 시계열이 아닌 스칼라 시계열을 전제하기로 한다. 시계열은 다음과 같이 내재된 모형을 따라서 전계된다는 가정을 한다.

$$x_t = Ax_{t-1} + \omega_t$$

→ x_t는 모형이 제시하는 값이며 직접 측정할 수는 없다.

→ A는 “변화 계수” (transition coefficient) 이다.

→ A가 1에 가까우면 random walk이고 A가 0에 가까우면 white noise이다.

Fast campus
Copyright FASTCAMPUS Corp. All Rights Reserved

Kalman 필터: 개요

- 여기에서는 벡터 시계열이 아닌 스칼라 시계열을 전제하기로 한다. 시계열은 다음과 같이 내재된 모형을 따라서 전계된다는 가정을 한다.

$$x_t = Ax_{t-1} + \omega_t$$

→ ω_t는 “프로세스 노이즈” (process noise)라고 불리우고 평균이 0이고 분산이 Q인 정규확률분포를 따른다고 가정한다: $\omega_t \sim N(0, Q)$.

→ Q는 “프로세스 분산” (process variance)이라고 불리운다.

Copyright FASTCAMPUS Corp. All Rights Reserved

I Kalman 필터: 개요

- 내재된 모형과 실제 데이터와의 연결은 다음 수식이 담당한다.

$$z_t = H\, x_t + \nu_t$$

→ z_t는 실제로 측정된 시계열 데이터 이다.

→ H는 “변환 계수” (transformation coefficient)이다.

→ ν_t는 “측정 노이즈” (measurement noise)라고 불리우고 평균은 0이고 분산은 R인 정규확률분포를 따른다고 가정한다: $\nu_t \sim N(0, R)$.

→ R은 “측정 분산” (measurement variance)이라고 불리운다.

Fast campus
Copyright FASTCAMPUS Corp. All Rights Reserved

I Kalman 필터: 알고리즘

- 칼만 필터는 반복적 알고리즘으로 구현할 수 있다.

→ 초기화 상태에서 예측 스텝과 업데이트 스텝을 반복하며 나간다.

1). 예측 스텝 (prediction step)은 다음 수식으로 요약할 수 있다.

$$\hat{x}_{t|t-1} = A\,\hat{x}_{t-1|t-1}$$

$$p_{t|t-1} = A^2\,p_{t-1|t-1} + Q$$

→ $\hat{x}_{t|t-1}$와 $p_{t|t-1}$같이 서브인덱스가 "$t|t-1$"이면 예측값이다.

→ $\hat{x}_{t-1|t-1}$와 $p_{t-1|t-1}$같이 서브인덱스가 "$t-1|t-1$"이면 업데이트 확정된 값이다.

Copyright FASTCAMPUS Corp. All Rights Reserved

I Kalman 필터: 알고리즘

2). 업데이트 스텝 (update step)은 다음 수식으로 요약할 수 있다.

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + k_t(z_t - H\hat{x}_{t|t-1})$$

$$p_{t|t} = p_{t|t-1}(1 - k_t H)$$

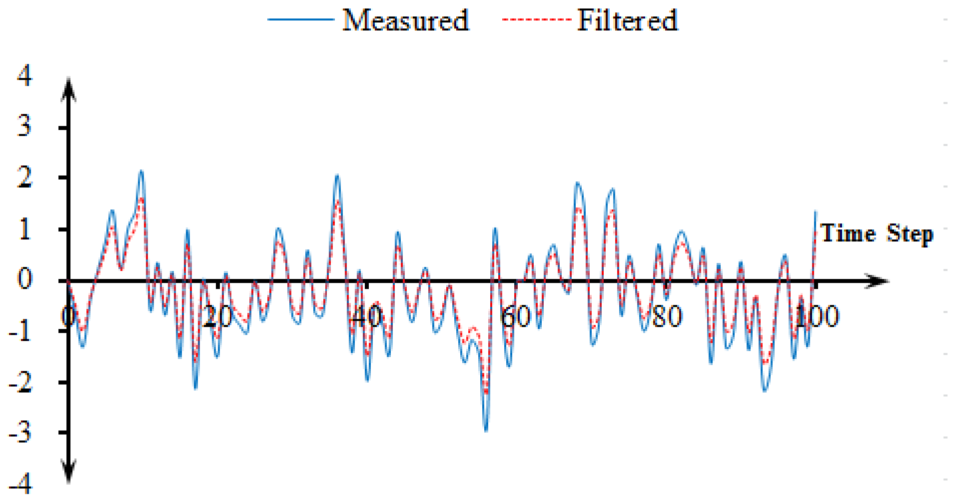
$$k_t = \frac{p_{t|t-1}H}{H^2\, p_{t|t-1} + R}$$

→ $\hat{x}_{t|t}$와 $p_{t|t}$같이 서브인덱스가 “$t|t$”이면, 이것도 업데이트가 확정된 값을 의미한다.

→ k_t는 “Kalman gain”이라 불리운다.

Copyright FASTCAMPUS Corp. All Rights Reserved

I Kalman 필터: 예



→ $A = 0.1$이며 내재모형은 white noise에 가까운 시계열이다.

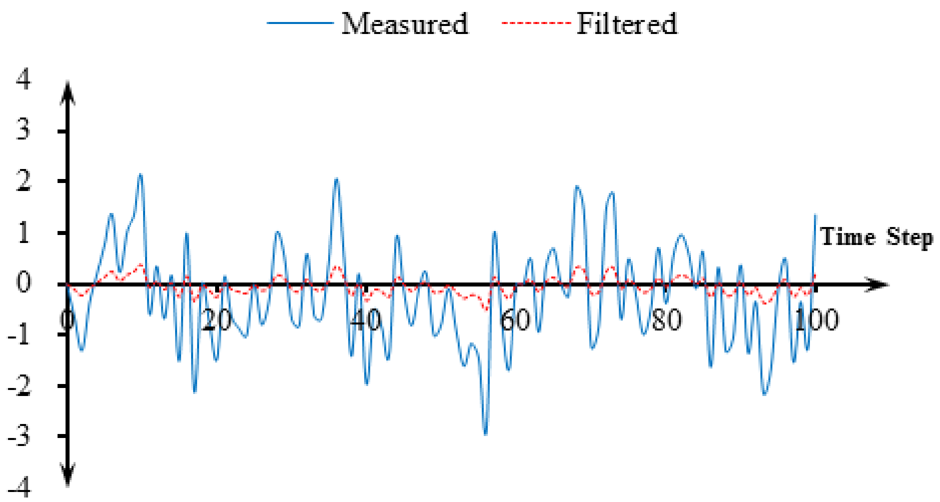
→ 프로세스 분산 $Q = 3$, 측정 분산 $R = 1$이다.

→ 이와같이 $Q > R$이면 필터된 값은 측정값에 가깝다.

Copyright FASTCAMPUS Corp. All Rights Reserved

Kalman 필터: 예

Life Changing Education



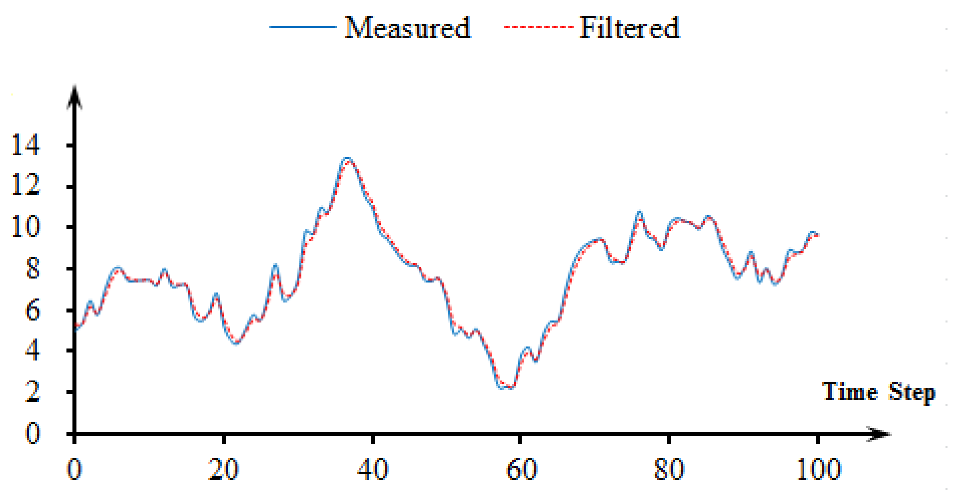
→ $A = 0.1$이며 내재모형은 white noise에 가까운 시계열이다.

→ 프로세스 분산 $Q = 0.2$, 측정 분산 $R = 1$이다.

→ 이와같이 $Q < R$이면 필터된 값은 내재모형에 가깝다.

Copyright FASTCAMPUS Corp. All Rights Reserved

I Kalman 필터: 예



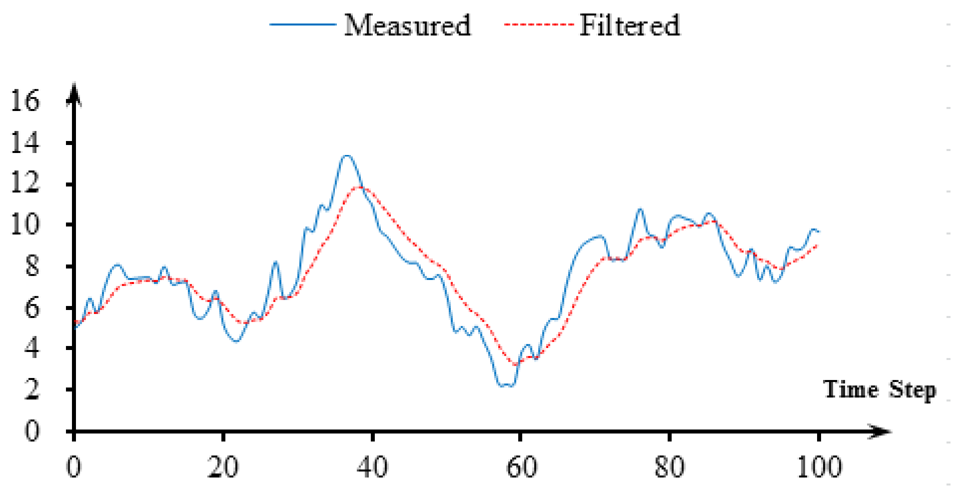
→ $A = 1$이며 내재모형은 random walk 시계열이다.

→ 프로세스 분산 $Q = 2$, 측정 분산 $R = 1$이다.

→ 이와같이 $Q > R$이면 필터된 값은 측정값에 가깝다.

Life Changing Education
Copyright FASTCAMPUS Corp. All Rights Reserved

I Kalman 필터: 예



→ $A = 1$이며 내재모형은 random walk 시계열이다.

→ 프로세스 분산 $Q = 0.1$, 측정 분산 $R = 1$이다.

→ 이와같이 $Q < R$이면 필터된 값은 내재모형에 가깝다.

Copyright FASTCAMPUS Corp. All Rights Reserved

| 끝.

Life Changing Education

감사합니다.

Fast campus

Fast campus

Copyright FASTCAMPUS Corp. All Rights Reserved