Bayesian Update Methods for Effect-based Prognostics

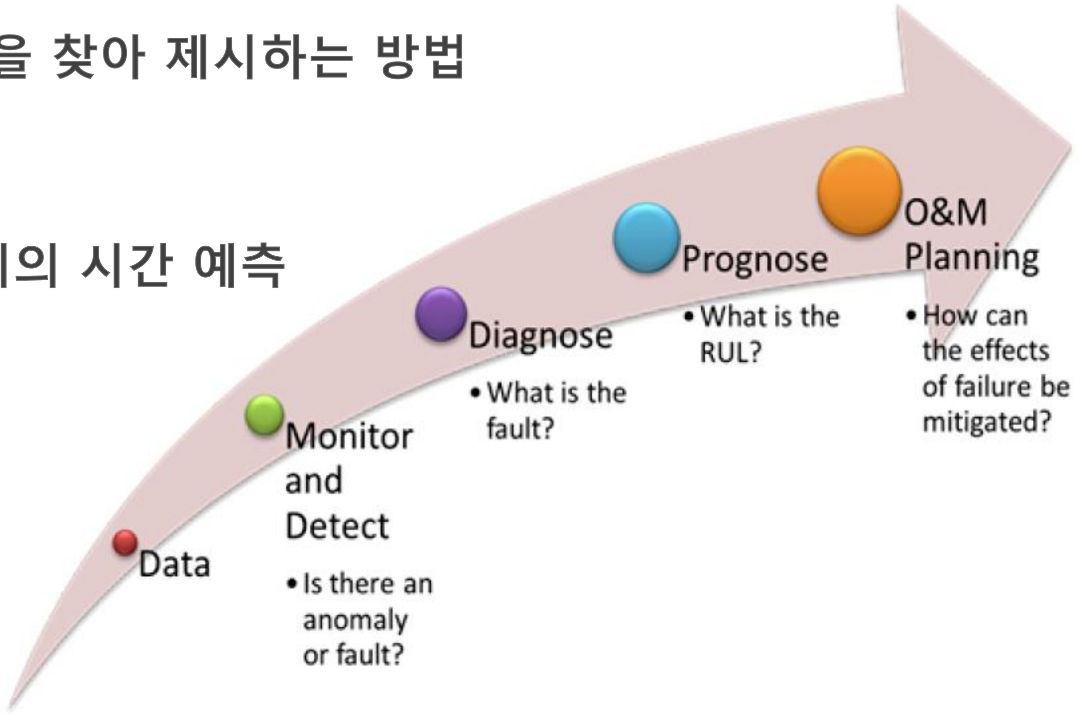
2018 추계 KNS 워크샵

M - 원전 상태감시 및 진단 신기술과 적용

경희대학교 허 균 영

Introduction

- Monitoring
  - 운전상태와 관련된 신호 및 특징을 감시 현재값과 예측값을 비교
  - Where we should be vs Where we are
- Diagnostic
  - 이상(Anomaly)의 근본 원인을 찾아 제시하는 방법
  - 분류(Classification)의 문제
- Prognostic
  - 고장(Failure)로 진전되기까지의 시간 예측



- Prognostics의 특징
  - 신뢰도 공학과 다른 점
    - 신뢰도 공학은 주로 과거 Generic Data에 기반하여 잔여 수명을 예측하는 것으로 설계 단계에서 중요시 다룸
  - 진단 공학과 다른 점
    - 진단 공학은 주로 현재 측정 Data에 기반하여 고장의 종류를 판단하는 것으로서 일반적으로 잔여 수명을 예측하지는 않음

- 따라서 현재 측정 Data를 반영하여 잔여 수명을 예측한다는 점에서 기존 학문 분야와 차이가 남

Prognostic (가용한 데이터에 따른 분류)

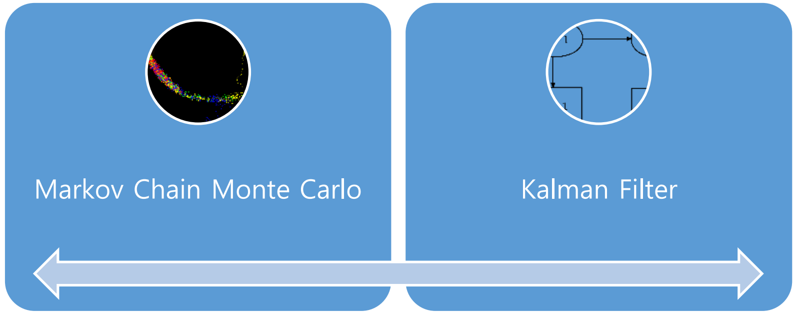
- Type I (Reliability based)
  - 전통적인 신뢰도분석, 평균적인 기기/시스템의 평균적인 운전조건을 고려 유효수명 예측
- Type II (Stressor based)
  - 운전조건을 고려하여 예측, 평균적인 기기/시스템의 특정 운전조건을 고려 유효수명 예측
- Type III (Effect based)
  - 운전조건 및 경년열화를 고려하여 예측, 특정 기기/시스템의 특정 운전조건을 고려 유효수명 예측

Prognostic (모델링 방법에 따른 분류)

- 고장예지 방법은 모델의 사용 여부에 따라, 데이터 기반 방법과 모델 기반 방법으로 분류됨

데이터 기반 방법 :

정형화된 모델없이, 과거의 고장 데이터로부터 추출된 열화 특성을 미래 시점으로 외삽하는 기술



모델 기반 방법 :

정형화된 모델의 모수에 과거의 고장 데이터로부터 추출된 정보를 입력하고, 그 모델을 이용하여 열화 특성을 외삽하는 기술

Kalman Filter

- Optimal Bayesian
- 시스템 모델 및 측정이 선형으로 표현
- 모델 및 측정 오차가 가우시안 분포를 따른다고 가정

Grid-based Method

- Optimal filter for Discrete State Space
- Hidden Markov Model으로도 불리고, 분포 및 적분을 차분화하여 계산 가능

Extended KF

- Approximated Bayesian
- 시스템 모델 및 측정이 비선형인 경우 Taylor series를 이용하여 선형으로 근사화

Unscented KF

- Approximated Bayesian
- 시스템 모델 및 측정이 비선형인 경우 가우스 커널을 중첩하여 근사화

Particle Filter

- 몬테카를로 시뮬레이션을 통한 Recursive 베이즈 필터
- Sequential Importance Sampling과 Resampling 방법에 따라 다양하게 세부 분류

예시문제

– **Data (기존에 수집된 열화 패턴)**

x	1	2	3	4	5
y	1.8	6.3	14	24	36.3

$\sigma_{y_i}^2 = 1$

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \epsilon_i \qquad \epsilon \sim \mathrm{N}(0, \sigma^2)$$

$$\beta_{prior} = \begin{bmatrix} 1.502 \\ 0.517 \\ 0.938 \end{bmatrix} \qquad \sigma_{\beta_{prior}}^2 = \begin{bmatrix} 13.088 \\ 2.283 \\ 0.018 \end{bmatrix}$$

2차 다항식으로 모델 구성

Method – Kalman Filter

- **칼만 필터** : 시스템 모델을 이용한 상태 예측과 측정값 업데이트를 반복적으로 수행
  - 순차적으로 들어오는 관측데이터를 이용해 대상의 상태에 대해 연속적인 베이지안 업데이트를 수행
  - 시스템 모델과 측정 모델에 대해 Gaussian Noise와 Linear 모델을 가정
  - Bayesian Filter 개념 : 시스템 모델 $f(y_{i-1}, \Theta)$를 이용하여 상태 y_i를 예측하고, 예측값을 관측값 z_i로 보정

$$y_i = Ay_{i-1} + B + w_i \qquad w_i \sim \mathrm{N}(0, Q_i)$$

$$z_i = Hy_i + v_i \qquad w_i \sim \mathrm{N}(0, R_i)$$

Method

① **Prediction: 이전 스텝의 상태를 가지고, 현재 스텝의 상태를 예측하는 것**

$$\bar{\mu}_{y_i} = A\mu_{y_{i-1}} + B$$

$$\bar{Q}_i = AQ_{i-1}A^T$$

Method

② **Update: 지금 스텝의 관측값을 가지고 지금 스텝의 예측된 상태를 업데이트 하는 것**

- **Kalman Gain**

$$K_i = \bar{Q}_i H^T (H\bar{Q}_i H^T + R_i)^{-1}$$

- **Measurement update**

$$\mu_{y_i} = \bar{\mu}_{y_i} + K_i \left(z_i - H\bar{\mu}_{y_i}\right)$$
$$Q_i = (I - K_i H)\bar{Q}_i$$

Method

- **Bayesian update – Gaussian distribution**

$$N(\mu_1, \sigma_1^2) : prior$$
$$N(\mu_2, \sigma_2^2) : Likelihood$$

$$\mu_1' = \frac{\mu_1\sigma_2^2 + \mu_2\sigma_1^2}{\sigma_1^2 + \sigma_2^2}$$
$$= \mu_1 - \mu_1\frac{\sigma_1^2 + \sigma_2^2}{\sigma_1^2 + \sigma_2^2} + \frac{\mu_1\sigma_2^2 + \mu_2\sigma_1^2}{\sigma_1^2 + \sigma_2^2}$$
$$= \mu_1 + \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}(\mu_2 - \mu_1)$$

$$\sigma_1'^2 = \frac{\sigma_1^2\sigma_2^2}{\sigma_1^2 + \sigma_2^2}$$
$$= \sigma_1^2 - \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}\sigma_1^2$$

$\mu_1 \rightarrow H\mu_1$

$\sigma_1 \rightarrow H\sigma_1$

Kalman gain

$$\mu_1' = \mu_1 + \boxed{\frac{H\sigma_1^2}{H^2\sigma_1^2 + \sigma_2^2}}(\mu_2 - H\mu_1)$$

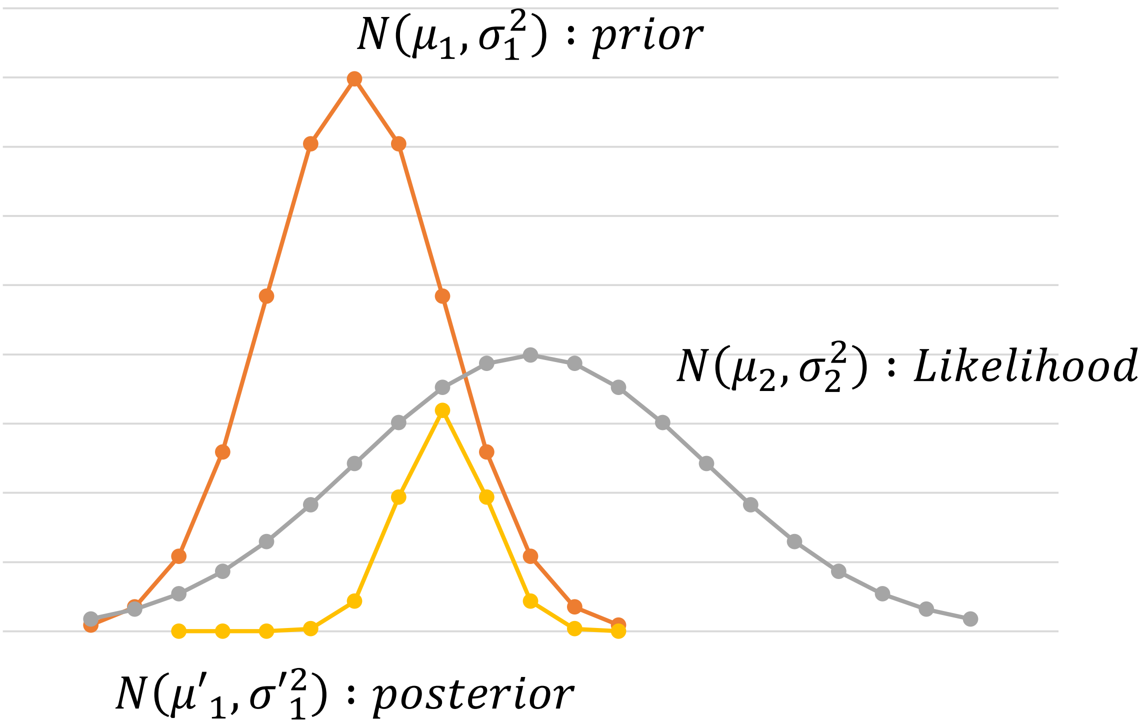
$$\sigma_1'^2 = \sigma_1^2 - \boxed{\frac{H\sigma_1^2}{H^2\sigma_1^2 + \sigma_2^2}}H\sigma_1^2$$

$$K_i = \bar{Q}_i H^T (H\bar{Q}_i H^T + R_i)^{-1}$$

prediction error

$$\mu_{y_i} = \boxed{\bar{\mu}_{y_i}} + K_i \boxed{\left(z_i - H\bar{\mu}_{y_i}\right)}$$

$$Q_i = (I - K_i H)\bar{Q}_i$$


$N(\mu_1, \sigma_1^2)$: $prior$
$N(\mu_2, \sigma_2^2)$: $Likelihood$
$N(\mu'_1, \sigma'^2_1)$: $posterior$

Method

② **Update: 지금 스텝의 관측값을 가지고 지금 스텝의 예측된 상태를 업데이트 하는 것**

- **Measurement uncertainty**

$$K_i = \bar{Q}_i H^T (H\bar{Q}_i H^T + R_i)^{-1}$$

Low uncertainty	High uncertainty
$R \to 0$	$R \to \infty$
$K_i = \bar{Q}_i H^T (H\bar{Q}_i H^T + 0)^{-1}$ $= \bar{Q}_i H^T H^{T^{-1}} \bar{Q}_i^{-1} H^{-1}$ $= H^{-1}$ $\mu_{y_i} = \bar{\mu}_{y_i} + K_i \left(z_i - H\bar{\mu}_{y_i}\right)$ $= \bar{\mu}_{y_i} + H^{-1} z_i - H^{-1} H \bar{\mu}_{y_i}$ $= \boxed{H^{-1} z_i}$ ➔ 측정값 $Q_i = (I - K_i H)\bar{Q}_i$ $= 0$	$K_i = \bar{Q}_i H^T (H\bar{Q}_i H^T + \infty)^{-1}$ $= \bar{Q}_i H (\infty)^{-1}$ $= 0$ $\mu_{y_i} = \bar{\mu}_{y_i} + K_i \left(z_i - H\bar{\mu}_{y_i}\right)$ $= \boxed{\bar{\mu}_{y_i}}$ ➔ 예측값 $Q_i = (I - K_i H)\bar{Q}_i$ $= \bar{Q}_i$

Example 1

– **Data**

x	1	2	3	4	5
y	1.8	6.3	14	24	36.3

$\sigma_{y_i}^2 = 1$

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \epsilon_i \qquad \epsilon \sim \mathrm{N}(0, \sigma^2)$$

$$\beta_{prior} = \begin{bmatrix} 1.502 \\ 0.517 \\ 0.938 \end{bmatrix} \quad \sigma_{\beta_{prior}}^2 = \begin{bmatrix} 13.088 \\ 2.283 \\ 0.018 \end{bmatrix}$$

– **Model**

- **State**

$$\beta_i = A\beta_{i-1} + B + w_i, \qquad w_i \sim N(0, Q_i), \qquad \beta_i = [\beta_{0,i}, \beta_{1,i}, \beta_{2,i}]^T, A = I_3, B = \emptyset, \quad Q_i = \begin{bmatrix} \sigma_{\beta_{0,i}}^2 & 0 & 0 \\ 0 & \sigma_{\beta_{1,i}}^2 & 0 \\ 0 & 0 & \sigma_{\beta_{2,i}}^2 \end{bmatrix}$$

- **Measurement**

$$y_i = H_i\beta_i + v_i, \qquad v_i \sim \mathrm{N}(0, R_i), \qquad (H_i = [1, i, i^2]\ (\because x_i = i), R_i = \sigma_{y_i}^2)$$

Example

① **Prediction**

$$\boxed{\begin{aligned}\bar{\mu}_{\beta_i} &= A\mu_{\beta_{i-1}} \\ \bar{Q}_i &= AQ_{i-1}A^T\end{aligned}} \Leftarrow \quad A = \begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix}, \quad \mu_{\beta_i} = \begin{bmatrix}\mu_{\beta_{0,i}}\\ \mu_{\beta_{1,i}}\\ \mu_{\beta_{2,i}}\end{bmatrix}, \quad Q_i = \begin{bmatrix}\sigma^2_{\beta_{0,i}} & 0 & 0\\ 0 & \sigma^2_{\beta_{1,i}} & 0\\ 0 & 0 & \sigma^2_{\beta_{2,i}}\end{bmatrix}$$

② **Update**

- Kalman Gain

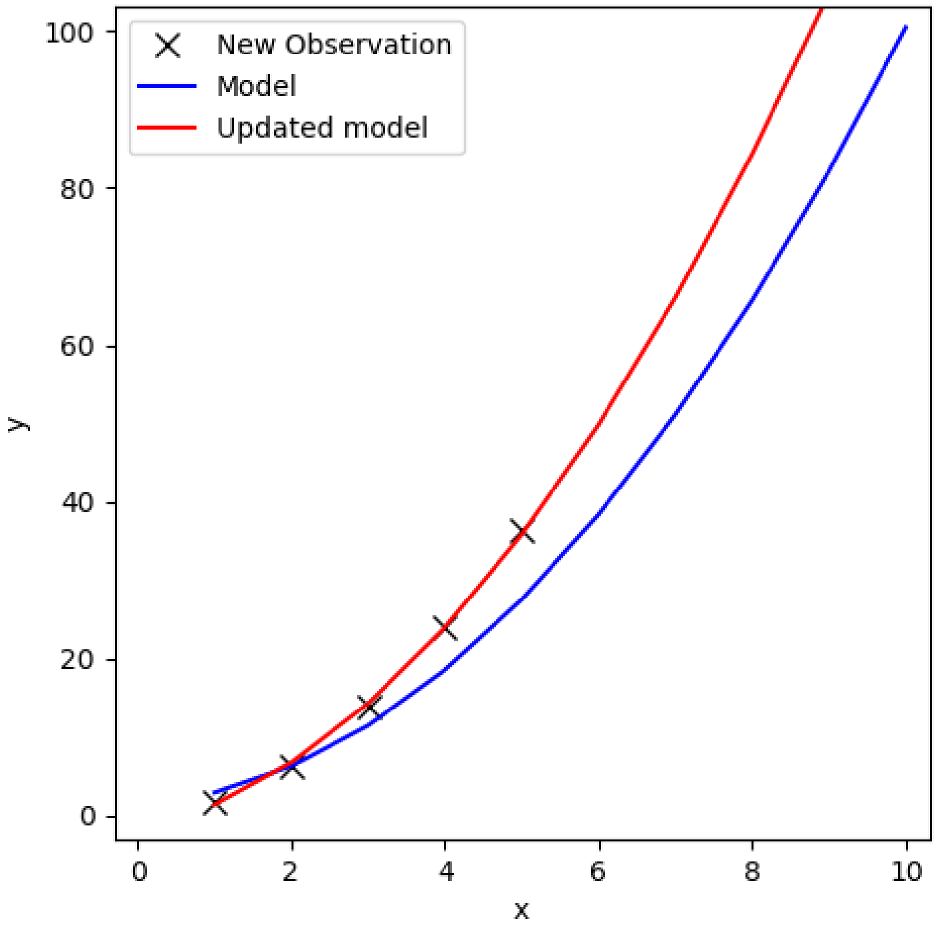
$$\boxed{K_i = \bar{Q}_i H^T (H\bar{Q}_i H^T + R_i)^{-1}} \Leftarrow \quad H_i = [1, i, i^2], \quad R_i = \sigma^2_{y_i}$$

- Measurement update

$$\boxed{\begin{aligned}\mu_{\beta_i} &= \bar{\mu}_{\beta_i} + K_i\left(y_i - H\bar{\mu}_{\beta_i}\right)\\ Q_i &= (I - K_i H)\bar{Q}_i\end{aligned}}$$

Example

– Result

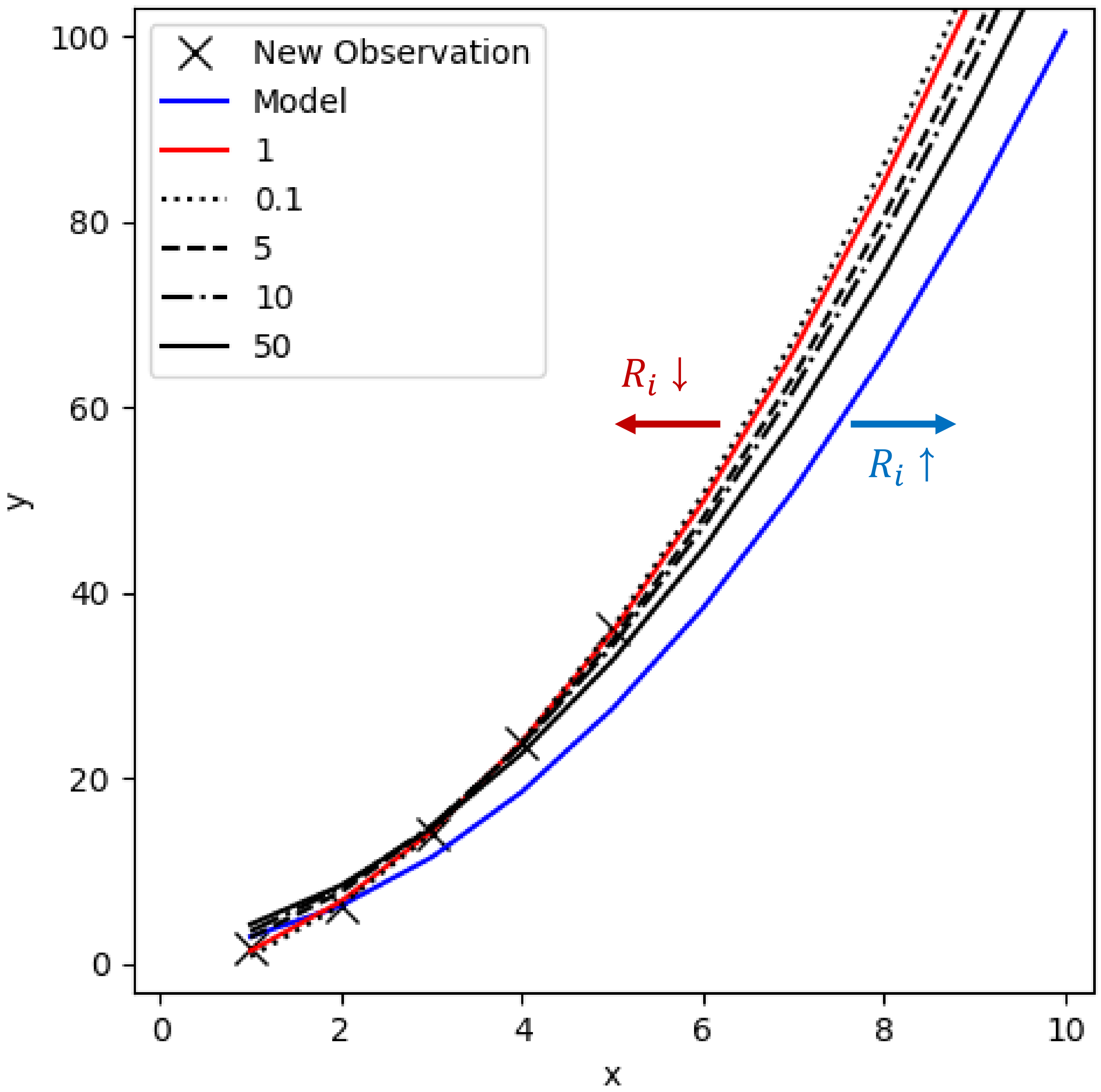


Example

$$K_i = \bar{Q}_i H^T (\boxed{H\bar{Q}_i H^T} + \boxed{R_i})^{-1}$$

Prior Information — New Observation

➔ Adjust R_i from 0.1~50



Method – Particle Filter

- **파티클 필터** : 베이지안 업데이트 방법을 이용하여 모델 모수에 대해 관측 정보를 "순차적"으로 업데이트
  - 순차적으로 들어오는 관측데이터를 이용해 모델에 대해 연속적인 베이지안 업데이트를 수행
  - 대량의 난수를 이용하여 Non-Gaussian Noise, Non-Linear 모델에 대해서 적용 가능
  - Bayesian Filter 개념 : 시스템 모델 $f(y_{i-1}, \Theta)$를 이용하여 상태 y_i를 예측하고, 예측값을 관측값 z_i로 보정

$$y_i = f(y_{i-1}, \Theta_i)$$

$$\begin{aligned} z_i &= h(y_i) + \epsilon \\ &= y_i + \epsilon, \qquad \epsilon \sim \mathrm{N}(0, \sigma^2) \end{aligned}$$

Method

① **Prediction: 이전 스텝의 상태를 가지고, 현재 스텝의 상태를 예측하는 것**

$$\Theta_i^j \sim p(\Theta_i|\Theta_{i-1}) = p(\Theta_{i-1}), \qquad j = 1, 2, \cdots n_s$$

$$y_{i-1}^j \sim p(y_{i-1}), \qquad j = 1, 2, \cdots n_s$$

$$y_i^j = f\left(y_{i-1}^j, \Theta_i^j\right), \qquad j = 1, 2, \cdots n_s$$

➔ y_i의 사전 분포 $p(y_i)$가 얻어짐

현 step 모델 파라미터 Θ_i^j 샘플링 – 모델 파라미터는 step에 영향을 받지 않으므로 이전 step의 파라미터 분포로부터 현 step 파라미터 분포 샘플링

이전 step 상태 y_{i-1}^j 샘플링 – 이전 step 상태의 분포 $p(y_{i-1})$로부터 샘플링

이전 step 상태 y_{i-1}^j와 현 step 모델 파라미터 Θ_i^j를 이용하여 시스템 모델로 부터 현 스텝 상태 y_i^j 계산

Method

- 파티클 필터

② **Update: 지금 스텝의 관측값을 가지고 지금 스텝의 상태를 업데이트 하는 것 (사후분포)**

- Monte Carlo Integration

$$E[g(x)] = \int g(x)p(x)dx\,, \qquad p(x) : PDF$$

1) $X^j \sim p(x),\ j = 1, 2, \cdots n_s$
2) $Y^j = g(X^j),\ j = 1, 2, \cdots n_s$
3) $E[g(x)] = \bar{Y} = \frac{1}{N}\sum_{j=1}^{n_s} Y^j$

Method

- 파티클 필터

② Update

- **Importance Sampling: 대상 분포가 알려지지 않아 직접 샘플링을 수행할 수 없는 경우, 알고 있는 분포 (q(x), importance distribution)를 도입하여, 우리가 잘 알고 있는 분포로부터 샘플링을 수행하여 기댓값을 구할 수 있음**

$$\int g(x)p(x)dx = \int g(x)\frac{p(x)}{q(x)}q(x)dx$$

$$= \int g(x)w(x)q(x)dx$$

* $q(x)$: Importance distribution

1) $X^j \sim q(x),\ j = 1, 2, \cdots n_s$
2) $Y^j = g(X^j)w(X^j),\ j = 1, 2, \cdots n_s$
3) $E[g(x)] = \bar{Y} = \frac{1}{N}\sum_{j=1}^{n_s} Y^j$

Note: 여전히 p(x)를 구해야 하는 현안이 있음

Method

- **파티클 필터**

② **Update**

- Update 단계에서는 관측값 z_i를 이용하여 y_i의 사후분포를 구하고자 함

$$p(y_i|z_i) = ? \qquad \frac{p(y_i|z_i)}{q(y_i)} q(y_i) = \frac{p(z_i|y_i)p(y_i)}{q(y_i)} q(y_i)$$

→ Importance distribution $\boldsymbol{q}(y_i)$를 사전 분포 $p(y_i)$로 두면,

$$w(y_i) = \frac{p(z_i|y_i)p(y_i)}{p(y_i)} = p(z_i|y_i) \Rightarrow Likelihood$$

→ 정규분포일 경우, Likelihood 함수

$$w(y_i^j) = \mathrm{L}(z_i|y_i^j) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\left(z_i - y_i^j\right)^2}{2\sigma^2}\right), \qquad j = 1, 2, \cdots n_s$$

Method

- **파티클 필터**

③ **Resampling**

- Normalized weight: 분자/분모에 모두 weight를 배치함으로써 사후분포를 계산할 때 분모에 들어가는 적분이 서로 상쇄되어 없앨 수 있음

$$w(y_i^j) = \frac{w(y_i^j)}{\sum_{j=1}^{n_s} w(y_i^j)}$$

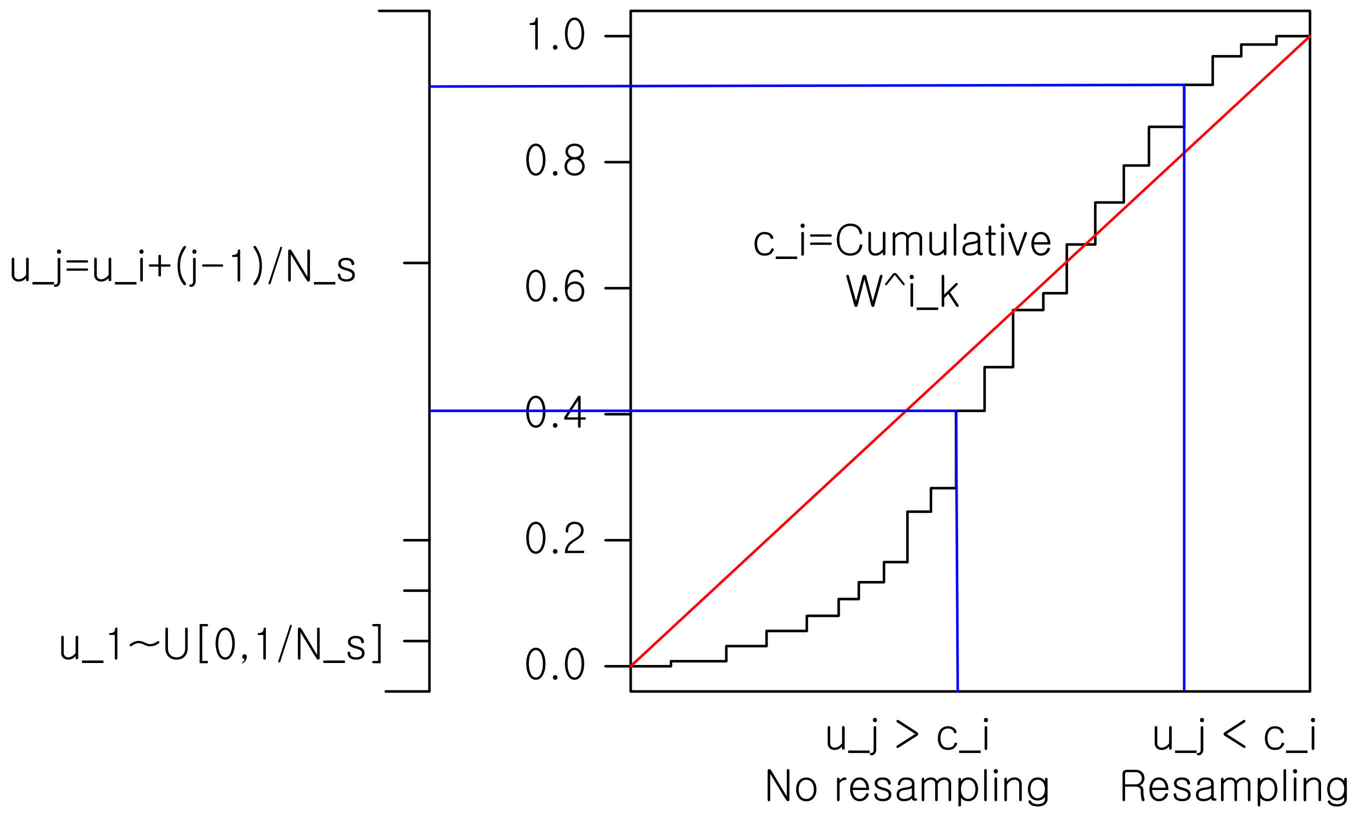
- Normalized weight의 누적분포함수로부터 Inverse transform sampling 기법을 이용하여 Resampling 수행
  - 문제는 소위 degenerate라고 해서 weight의 variance가 계속 증가함. 그러면 얼마 후에는 weight 중에서 하나만 1이 되고 나머지는 모두 0이 되어, 가중치가 의미가 없어짐
- Update 단계에서 얻은 $w(y_i)$를 이용하여 Resampling 수행

Method

- 파티클 필터

  ③ **Resampling**

  - Inverse transform sampling



Method

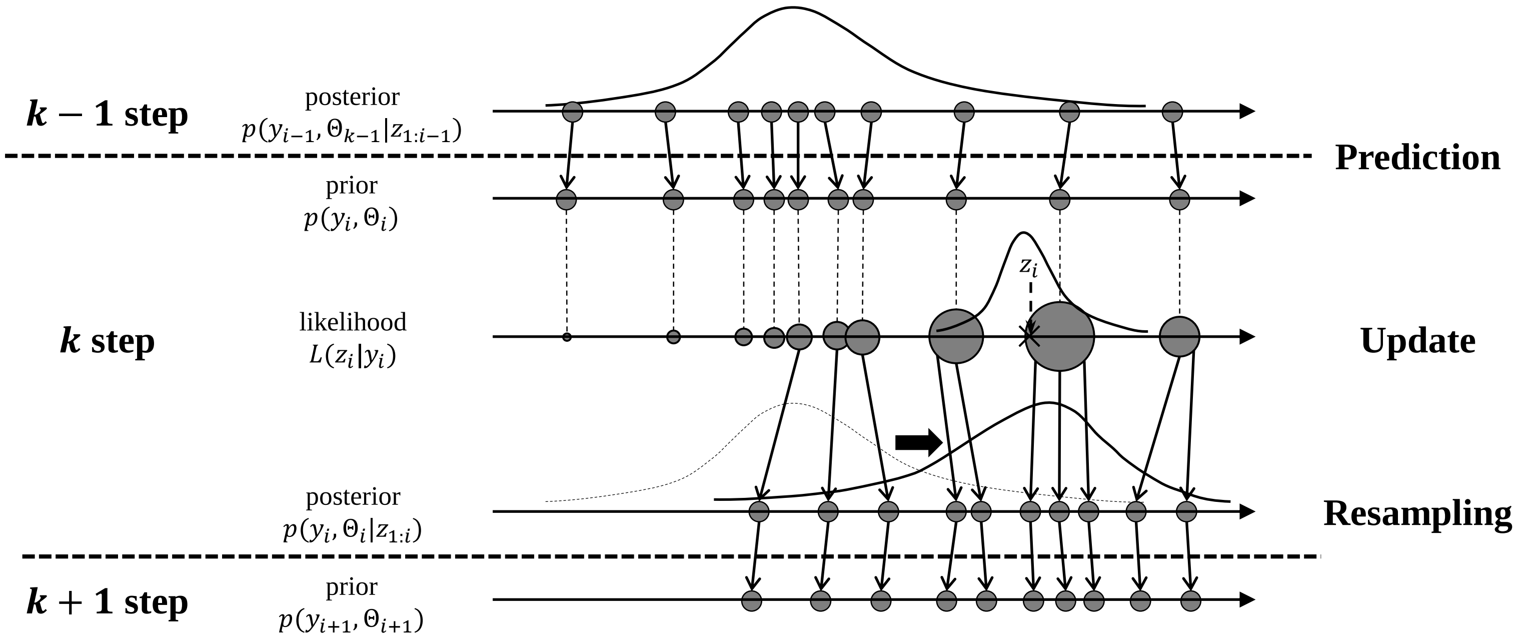
- **파티클 필터**

  ③ **Resampling**

  - y_i^j는 y_{i-1}^j와 Θ_i^j를 이용하여 계산됨 ← $y_i^j = f\left(y_{i-1}^j, \Theta_i^j\right)$
  - $w\left(y_i^j\right)$의 index j를 이용하여 Θ_i^j 또한 Resampling 수행
  - $p(y_i|z_i)$, $p(\boldsymbol{\Theta}_i|z_i)$ 사후 분포를 얻음

Method

- 파티클 필터



y_i : 예측된 상태
Θ_i : 모델 파라미터
z_i : 관측된 상태

Method

- **Balancing between Prior and Newly observed information**

$$w(y_i^j) = \mathrm{L}(z_i|y_i^j\ ) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{\left(z_i - y_i^j\right)^2}{2\sigma^2}\right), \qquad j = 1, 2, \cdots n_s$$

$$\lim_{\sigma\to\infty} L = \begin{cases} 0, & z_i = y_i^j \\ 0, & z_i \neq y_i^j. \end{cases}$$

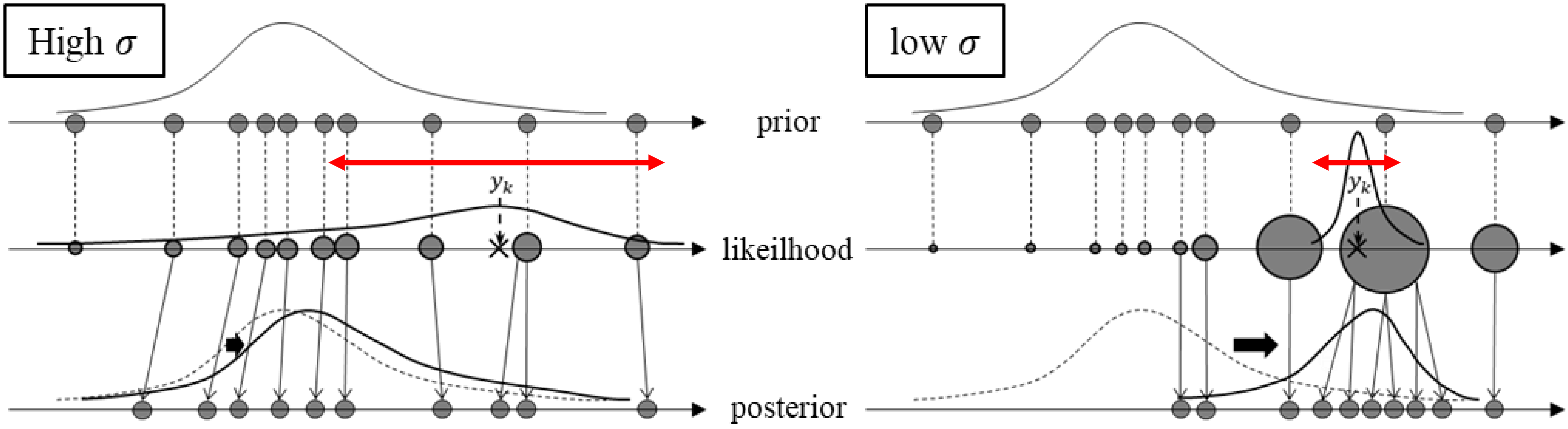
→ Likelihood 분포가 넓게 퍼짐
→ 모든 파티클들이 거의 균등하게 Resampling 됨
→ 따라서, 사전 분포 ≈ 사후 분포
→ 관측 정보가 덜 반영됨

$$\lim_{\sigma\to 0} L = \begin{cases} \infty, & z_i = y_i^j \\ 0, & z_i \neq y_i^j. \end{cases}$$

→ Likelihood 분포가 관측값에 가까운 파티클들에 집중됨
→ 관측값 주변의 파티클들이 집중적으로 Resampling 됨
→ 따라서, 사전 분포가 관측값 쪽으로 이동함
→ 관측 정보가 많이 반영됨

Method

- 관측 정보의 반영
  - 관측 정보의 불확실성에 따라 관측 정보를 어느 정도로 반영할 것인지 결정
  - 관측 정보의 불확실도 ↑ → 기존 모델을 더 신뢰 → 업데이트↓ → 기존 모델의 예측 결과를 그대로 사용
  - 관측 정보의 불확실도 ↓ → 관측 정보를 더 신뢰 → 업데이트↑ → 관측값 쪽으로 예측 분포 수정



Example 2

– **Data**

x	1	2	3	4	5
y	1.8	6.3	14	24	36.3

$\sigma_{y_i}^2 = 1$

$$\beta_{prior} = \begin{bmatrix} 1.502 \\ 0.517 \\ 0.938 \end{bmatrix} \quad \sigma_{\beta_{prior}}^2 = \begin{bmatrix} 13.088 \\ 2.283 \\ 0.018 \end{bmatrix}$$

– **Model**

- **State**

$$\beta_{ki} = f(\beta_{ki-1}) = \beta_{ki-1}, \qquad k = 0, 1, 2$$

- **Measurement**

$$\begin{aligned} y_i = h(\beta_{ki}) + \epsilon &= \beta_{0i} + \beta_{1i}x_i + \beta_{2i}x_i^2 + \epsilon \\ &= \beta_{0i} + \beta_{1i}i + \beta_{2i}i^2 + \epsilon, \qquad \epsilon \sim \mathrm{N}\left(0, \sigma_{y_i}^2\right)\ (\because x_i = i) \end{aligned}$$

Example

① **Prediction**

$$\beta_{ki-1}^{j} \sim p(\beta_{ki-1}), \qquad j = 1, 2, \cdots n_s$$
$$\beta_{ki}^{j} = f(\beta_{ki-1}^{j}), \qquad j = 1, 2, \cdots n_s \qquad \rightarrow p(\beta_{ki}) : prior$$

② **Update**

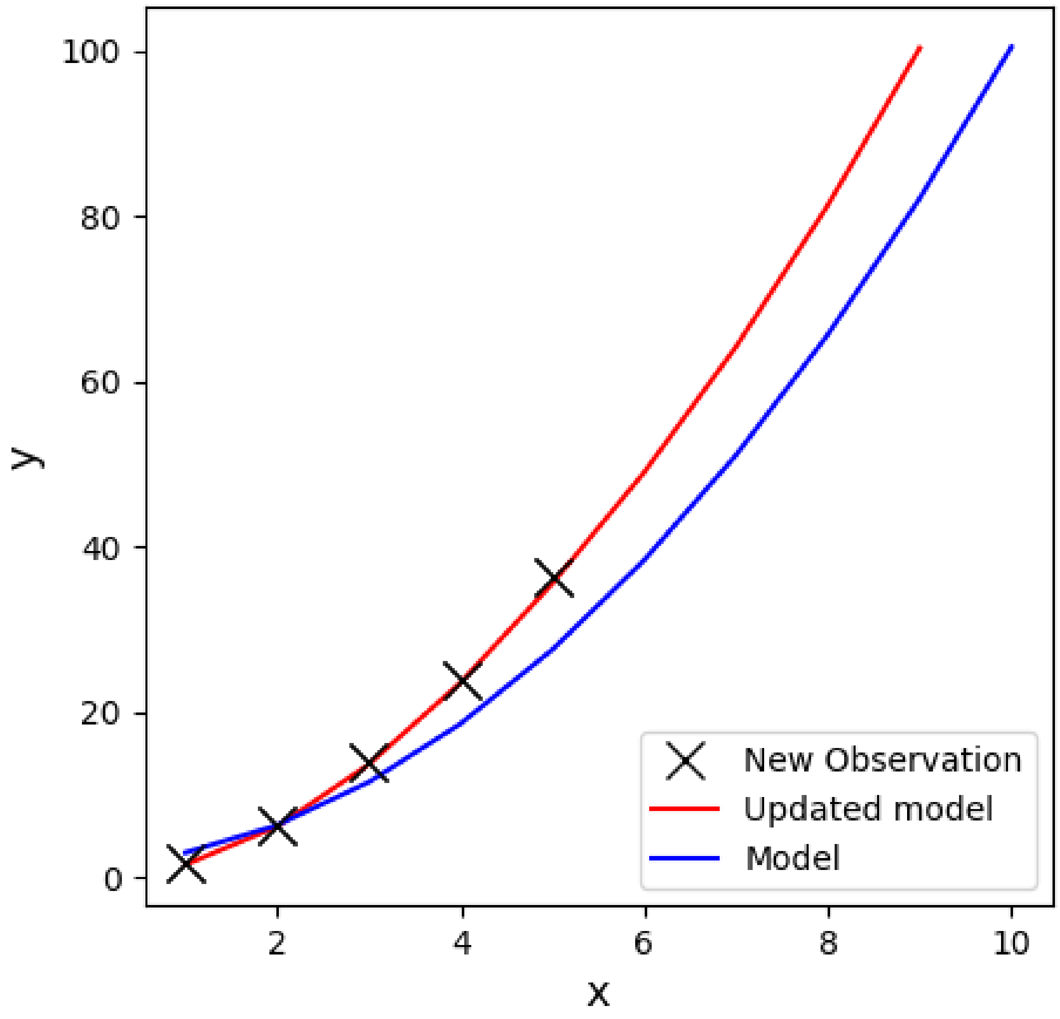
$$w(\beta_{ki}^{j}) = L(y_i | \beta_{ki}^{j}) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\left(y_i - h(\beta_{ki}^{j})\right)^2}{2\sigma^2}\right), \qquad j = 1, 2, \cdots n_s$$

$$w(\beta_{ki}^{j}) = \frac{w(\beta_{ki}^{j})}{\sum_{j=1}^{n_s} w(\beta_{ki}^{j})}, \qquad j = 1, 2, \cdots n_s$$

③ **Resampling**

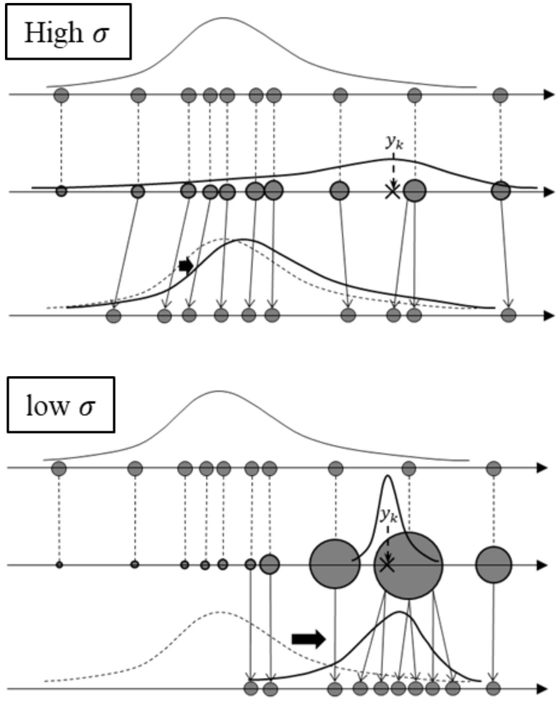
$$\beta_{ki}^{j} \sim w(\beta_{ki}), \qquad j = 1, 2, \cdots n_s \qquad \rightarrow p(\beta_{ki} | y_i) : posterior$$

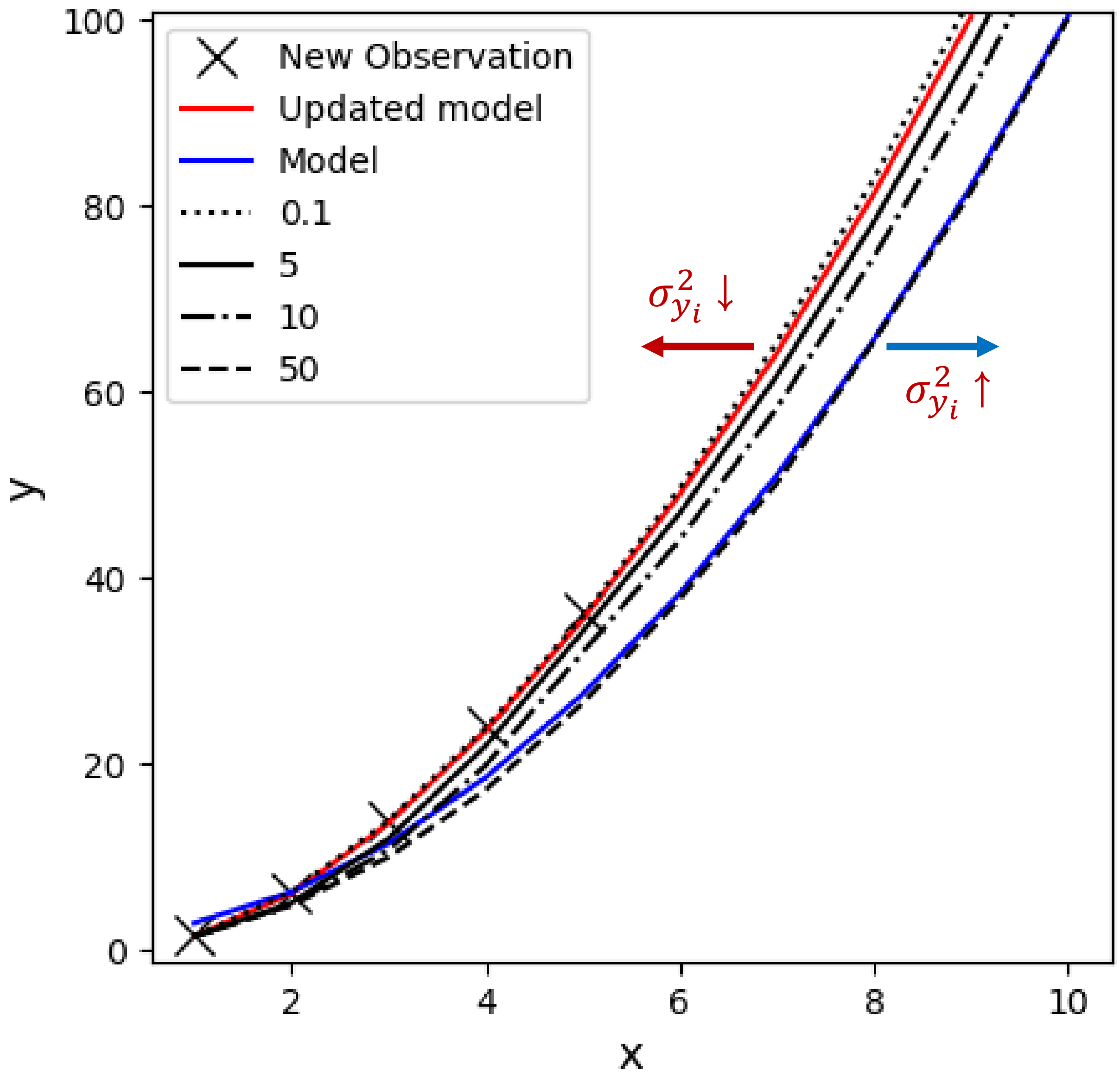
Example



100
80
60
40
20
0
y
2
4
6
8
10
x
New Observation
Updated model
Model

Example


High σ
y_k
low σ
y_k


New Observation
Updated model
Model
0.1
5
10
50
$\sigma^2_{y_i} \downarrow$
$\sigma^2_{y_i} \uparrow$
y
x
100
80
60
40
20
0
2
4
6
8
10

실례에서의 어려운 점

- 고장과 명시적인 상관성을 갖는 열화 현상을 찾을 수 있는지 여부
- 열화 정도가 측정가능한지 여부
- 측정 가능한 열화를 나타내는 기존 모델에서 채택하고 있는 계수의 uncertainty (standard deviation)과 현재 측정값의 uncertainty를 정량화하는 문제

EOD