Markov Chain Monte Carlo-based Manufacturing Process Control Algorithm

Kyungmin Kim, Jeyong Lee, and Minwoo Chae

Industrial Management and Engineering, Pohang University of Science and Technology

Motivation

- Achieving target properties of the final product is of importance in the manufacturing process.
- Manufacturing process control problem: The task of controlling particular process parameters to achieve predetermined target properties of the final product

Problem Statement

- The proposed algorithm assumes the following manufacturing process scenario:
  - There exist M sequential major manufacturing processes.
  - The manufacturing process is made until the m-th manufacturing process.
  - Among $(M-m)$ processes not made, the algorithm controls predetermined d_{control} process parameters to achieve K target properties of the final product.

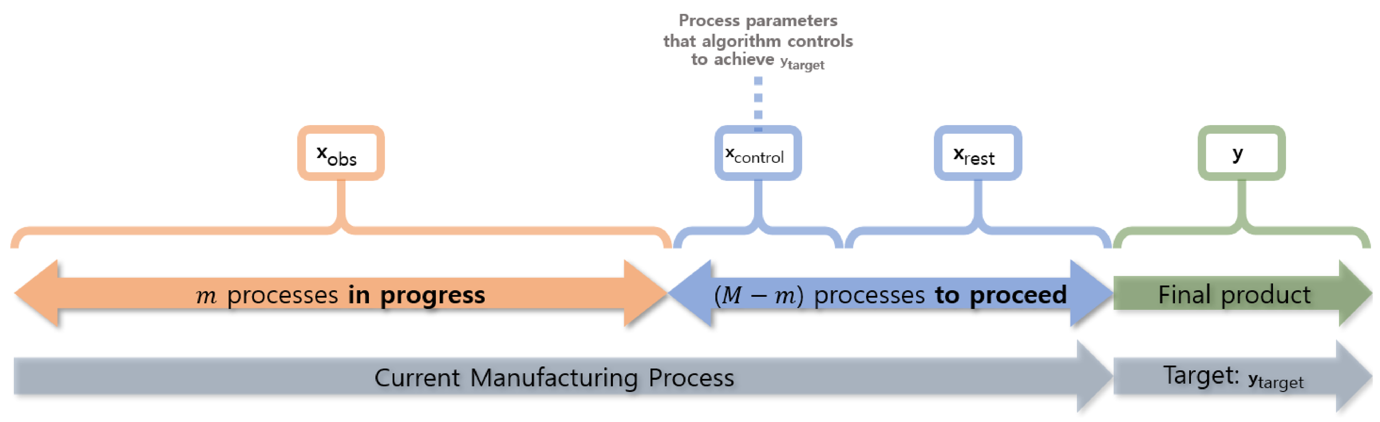

Figure 1:An Outline of the manufacturing process scenario

Notation

- $\mathbf{y}^i_{\text{target}}:=(y^i_{\text{target},1},\dots,y^i_{\text{target},K})$: Target properties of the i-th product
- $\mathbf{x}^i_{\text{mat}}:=(x^i_{\text{mat},1},\dots,x^i_{\text{mat},d_{\text{mat}}})$: The vector of proportions of raw materials to make the i-th product
  - d_{mat}: The number of raw materials used
- $\mathbf{x}^i_t:=(x^i_{t,1},\dots,x^i_{t,d_t})$: The vector of process parameters of the t-th manufacturing process
  - d_t: The number of process parameters of the t-th manufacturing processes
- For clarity, we subdivide the vector $\mathbf{x}:=(\mathbf{x}_{\text{mat}},\mathbf{x}_1,\dots,\mathbf{x}_M)$ into three vectors as follows. We have omitted the superscript $\boldsymbol{i}$ for brevity.
  - $\mathbf{x}_{\text{obs}}:=(x_{\text{obs},1},\dots,x_{\text{obs},d_{\text{obs}}})$, where $d_{\text{obs}}:=d_{\text{mat}}+\sum_{t=1}^{m} d_t$,
  - $\mathbf{x}_{\text{control}}:=(x_{\text{control},1},\dots,x_{\text{control},d_{\text{control}}})$
  - $\mathbf{x}_{\text{rest}}:=(x_{\text{rest},1},\dots,x_{\text{rest},d_{\text{rest}}})$, where $d_{\text{rest}}:=\sum_{t=m+1}^{M} d_t-d_{\text{control}}$

Modeling

- We model the relationship between $\mathbf{X}:=(\mathbf{X}_{\text{mat}},\mathbf{X}_1,\dots,\mathbf{X}_M)$ and $\mathbf{Y}:=\{Y_k\}_{k=1}^{K}$ as follows:

$$Y_k=f_k(\mathbf{X})+\epsilon_k, \tag{1}$$

  where $\mathbf{X}$ is i.i.d from the unknown distribution, $\epsilon_k\overset{\text{i.i.d}}{\sim}\mathcal{N}(0,\sigma_k^2)$, and ϵ_k is independent of $\mathbf{X}$.
- Re-expressing (1) using $\mathbf{X}_{\text{obs}},\mathbf{X}_{\text{control}}$, and $\mathbf{X}_{\text{rest}}$ is as follows:

$$Y_k=f_k(\mathbf{X}_{\text{obs}},\mathbf{X}_{\text{control}},\mathbf{X}_{\text{rest}})+\epsilon_k. \tag{2}$$

- The proposed algorithm is based on the probabilistic approach using the conditional distribution of $\mathbf{X}_{\text{control}}$ given $\mathbf{x}_{\text{obs}},\mathbf{x}_{\text{rest}}$, and $\mathbf{y}_{\text{target}}$, which is modeled as follows:

$$p(\mathbf{x}_{\text{control}}|\mathbf{x}_{\text{obs}},\mathbf{x}_{\text{rest}},\mathbf{y}_{\text{target}})\propto\prod_{k=1}^{K} p(y_{\text{target},k}|\mathbf{x}_{\text{obs}},\mathbf{x}_{\text{control}},\mathbf{x}_{\text{rest}}) \times p(\mathbf{x}_{\text{control}}|\mathbf{x}_{\text{obs}}). \tag{3}$$

  - The algorithm proposes $\mathbf{x}_{\text{control}}$ on the high density region of (3).
  - For $k\in[K]$, we model

$$Y_k|\mathbf{x}_{\text{obs}},\mathbf{x}_{\text{control}},\mathbf{x}_{\text{rest}}\sim\mathcal{N}(f_k(\mathbf{x}_{\text{obs}},\mathbf{x}_{\text{control}},\mathbf{x}_{\text{rest}}),\sigma_k^2) \tag{4}$$

  - We model $(\mathbf{X}_{\text{obs}},\mathbf{X}_{\text{control}})$ using Gaussian mixture distribution as follows:

$$p(\mathbf{x}_{\text{obs}},\mathbf{x}_{\text{control}})=\sum_{g=1}^{G}\phi_g\mathcal{N}\left(\begin{pmatrix}\mathbf{x}_{\text{obs}}\\ \mathbf{x}_{\text{control}}\end{pmatrix};\begin{pmatrix}\mu_{\text{obs},g}\\ \mu_{\text{control},g}\end{pmatrix},\begin{pmatrix}\boldsymbol{\Sigma}_{\text{obs},g} & \boldsymbol{\Sigma}_{\text{obs,control},g}\\ \boldsymbol{\Sigma}_{\text{control,obs},g} & \boldsymbol{\Sigma}_{\text{control},g}\end{pmatrix}\right), \tag{5}$$

    where $\phi_g=$(probability that $(\mathbf{X}_{\text{obs}},\mathbf{X}_{\text{control}})$ is in cluster g).
  - Then, $p(\mathbf{x}_{\text{control}}|\mathbf{x}_{\text{obs}})=\sum_{g=1}^{G}\pi_g\mathcal{N}(\mathbf{x}_{\text{control}};\mu_g,\boldsymbol{\Sigma}_g)$, where $\mu_g=\mu_{\text{control},g}+\boldsymbol{\Sigma}_{\text{control,obs},g}\boldsymbol{\Sigma}^{-1}_{\text{obs},g}(\mathbf{x}_{\text{obs}}-\mu_{\text{obs},g})$, $\boldsymbol{\Sigma}_g=\boldsymbol{\Sigma}_{\text{control},g}-\boldsymbol{\Sigma}_{\text{control,obs},g}\boldsymbol{\Sigma}^{-1}_{\text{obs},g}\boldsymbol{\Sigma}_{\text{obs,control},g}$, and $\pi_g=\frac{\phi_g\mathcal{N}(\mathbf{x}_{\text{obs}};\mu_{\text{obs},g},\boldsymbol{\Sigma}_{\text{obs},g})}{\sum_{g'=1}^{G}\phi'_g\mathcal{N}(\mathbf{x}_{\text{obs}};\mu_{\text{obs},g'},\boldsymbol{\Sigma}_{\text{obs},g'})}$.

MCMC-based Manufacturing Process Control Algorithm

- The proposed algorithm is in two-fold as follows:
  - Estimation of (3) up to normalizing constant

$$\begin{aligned}\hat{p}(\mathbf{x}_{\text{control}}|\mathbf{x}_{\text{obs}},\mathbf{x}_{\text{rest}},\mathbf{y}_{\text{target}})&\propto\prod_{k=1}^{K}\hat{p}(y_{\text{target},k}|\mathbf{x}_{\text{obs}},\mathbf{x}_{\text{control}},\mathbf{x}_{\text{rest}})\times\hat{p}(\mathbf{x}_{\text{control}}|\mathbf{x}_{\text{obs}})\\ &=\prod_{k=1}^{K}\mathcal{N}(y_{\text{target},k};\hat{f}_k(\mathbf{x}_{\text{obs}},\mathbf{x}_{\text{control}},\mathbf{x}_{\text{rest}}),\hat{\sigma}_k^2)\\ &\quad\times\sum_{g=1}^{G}\hat{\pi}_g\mathcal{N}(\mathbf{x}_{\text{control}};\hat{\mu}_g,\hat{\boldsymbol{\Sigma}}_g),\end{aligned} \tag{6}$$

    where $\hat{\mu}_g=\hat{\mu}_{\text{control},g}+\hat{\boldsymbol{\Sigma}}_{\text{control,obs},g}\hat{\boldsymbol{\Sigma}}^{-1}_{\text{obs},g}(\mathbf{x}_{\text{obs}}-\hat{\mu}_{\text{obs},g})$, $\hat{\boldsymbol{\Sigma}}_g=\hat{\boldsymbol{\Sigma}}_{\text{control},g}-\hat{\boldsymbol{\Sigma}}_{\text{control,obs},g}\hat{\boldsymbol{\Sigma}}^{-1}_{\text{obs},g}\hat{\boldsymbol{\Sigma}}_{\text{obs,control},g}$, and $\hat{\pi}_g=\frac{\hat{\phi}_g\mathcal{N}(\mathbf{x}_{\text{obs}};\hat{\mu}_{\text{obs},g},\hat{\boldsymbol{\Sigma}}_{\text{obs},g})}{\sum_{g'=1}^{G}\hat{\phi}_{g'}\mathcal{N}(\mathbf{x}_{\text{obs}};\hat{\mu}_{\text{obs},g'},\hat{\boldsymbol{\Sigma}}_{\text{obs},g'})}$ for $g\in[G]$.
    - Choose $\hat{f}_k(\cdot)$ with smallest MSE among considered models using K-fold cross-validation and take corresponding MSE as $\hat{\sigma}_k^2$ for $k\in[K]$.
    - Estimate $\mu_{\text{obs}_g},\mu_{\text{control}_g},\phi_g,\boldsymbol{\Sigma}_{\text{obs},g},\boldsymbol{\Sigma}_{\text{obs,control},g}$, and $\boldsymbol{\Sigma}_{\text{control},g}$, given the combination of covariance type and G that provides the smallest BIC score.
  - Proposition of $\hat{\mathbf{x}}_{\text{control}}$ based on Metropolis-Hastings algorithm
    - Assume $\mathbf{x}_{\text{obs}},\mathbf{x}_{\text{rest}}$, and $\mathbf{y}_{\text{target}}$ is given.
    - Step 1. Initialize $\mathbf{x}^{(0)}_{\text{control}}$.
    - Step 2. For $s=0$ to $n_{\text{samples}}-1$:
      (a) Generate $\tilde{\mathbf{x}}_{\text{control}}\sim q(\cdot|\mathbf{x}^{(s)}_{\text{control}})$, where $q(\tilde{\mathbf{x}}_{\text{control}}|\mathbf{x}^{(s)}_{\text{control}})=\mathcal{N}(\tilde{\mathbf{x}}_{\text{control}};\mathbf{x}^{(s)}_{\text{control}},\boldsymbol{\Sigma}_0)$.
      (b) Set $\mathbf{x}^{(s+1)}_{\text{control}}=\begin{cases}\tilde{\mathbf{x}}_{\text{control}} & \text{with prob. } \alpha(\mathbf{x}^{(s)}_{\text{control}},\tilde{\mathbf{x}}_{\text{control}})\\ \mathbf{x}^{(s)}_{\text{control}} & \text{with prob. } 1-\alpha(\mathbf{x}^{(s)}_{\text{control}},\tilde{\mathbf{x}}_{\text{control}}),\end{cases}$
      where $\alpha(\mathbf{x}^{(s)}_{\text{control}},\tilde{\mathbf{x}}_{\text{control}})=\min\{1,\frac{\prod_{k=1}^{K}\mathcal{N}(y_{\text{target},k};\hat{f}_k(\mathbf{x}_{\text{obs}},\tilde{\mathbf{x}}_{\text{control}},\mathbf{x}_{\text{rest}}),\hat{\sigma}_k^2)\times\sum_{g=1}^{G}\hat{\phi}_g\mathcal{N}(\tilde{\mathbf{x}}_{\text{control}};\hat{\mu}_g,\hat{\boldsymbol{\Sigma}}_g)}{\prod_{k=1}^{K}\mathcal{N}(y_{\text{target},k};\hat{f}_k(\mathbf{x}_{\text{obs}},\mathbf{x}^{(s)}_{\text{control}},\mathbf{x}_{\text{rest}}),\hat{\sigma}_k^2)\times\sum_{g=1}^{G}\hat{\phi}_g\mathcal{N}(\mathbf{x}^{(s)}_{\text{control}};\hat{\mu}_g,\hat{\boldsymbol{\Sigma}}_g)}\}$.
    - Step 3. Propose $\hat{\mathbf{x}}_{\text{control}}=(\frac{1}{n_{\text{samples}}}\sum_{s=1}^{n_{\text{samples}}}\mathbf{x}^{(s)}_{\text{control},i})_{i=1}^{d_{\text{control}}}$.

Evaluation Metric

- Unlike ordinary supervised learning settings, $\mathbf{y}$ is not given by the time the algorithm proposes $\hat{\mathbf{x}}_{\text{control}}$, which requires devising the proper evaluation metric.
- Two different evaluation metrics are considered, whose only difference lies in the way of setting $\mathbf{x}_{\text{rest}}$ as follows:
  - $\text{Score}_1:=|y_k-y_{\text{target},k}|-|\hat{f}_k(\mathbf{x}_{\text{obs}},\hat{\mathbf{x}}_{\text{control}},\mathbf{x}_{\text{rest}})-y_{\text{target},k}|$,
  - $\text{Score}_2:=|\hat{f}_k(\mathbf{x}_{\text{obs}},\mathbf{x}_{\text{control}},\bar{\mathbf{x}}_{\text{rest}})-y_{\text{target},k}|-|\hat{f}_k(\mathbf{x}_{\text{obs}},\hat{\mathbf{x}}_{\text{control}},\bar{\mathbf{x}}_{\text{rest}})-y_{\text{target},k}|$.
    - Assume validation set and corresponding target properties exist for evaluating process control algorithms.
    - The first metric sets $\mathbf{X}_{\text{rest}}$ as $\mathbf{x}_{\text{rest}}$, while the latter as $\bar{\mathbf{x}}_{\text{rest}}$, which is the predetermined target value.

Results

- Trace plot, autocorrelation plot, and effective sample size are used to check the convergence of the proposed algorithm.
- Used manufacturing data of $K=3,d_{\text{mat}}=15$, and $M=3$ $(d_1=12,d_2=4,d_3=10)$.
  - $m=3,d_{\text{control}}=3$
  - The validation set of size 116 is given.
- Due to the confidentiality issue, ratios of $\frac{|\hat{f}_k(\mathbf{x}_{\text{obs}},\mathbf{x}_{\text{control}},\bar{\mathbf{x}}_{\text{rest}})-y_{\text{target},k}|}{|\hat{f}_k(\mathbf{x}_{\text{obs}},\hat{\mathbf{x}}_{\text{control}},\bar{\mathbf{x}}_{\text{rest}})-y_{\text{target},k}|}$ are provided.

Figure 2:Ratios of discrepancies from $y_{\text{target},1}$ on a log scale.

Conclusion

- The proposed algorithm effectively controls process parameters to achieve the target properties of the final product in terms of the proposed evaluation metric.
- Dealing with heterogeneous data, a frequently occurring situation in the manufacturing process, needs to be addressed.
  - The evaluation metric depends on the accuracy of the predicted model.

contact E-mail:kkm959595@postech.ac.kr

[37673] 77, Cheongam-ro, Nam-gu, Pohang-si, Gyeongsangbuk-do, Korea