Robust Control Charts

Jayasankar Vattathoor
Graduate Student (MSc- Statistics)

Department of Mathematics & Statistics
Memorial University of Newfoundland, St. John’s

February 27, 2012
Outline

- Introduction to Control Charts
- Hotelling’s $T^2$ Control Chart
- Robust Control Charts
- Robust Estimators
- $T^2$ control chart based MCD/MVE estimators
Control Charts

- A graphical tool for detecting changes in the manufacturing conditions (shift in mean or variance) by comparing the actual observations with the control limits.

- Phase-I: Assessing the ability of the process - Estimation of process parameters and control limits

- Phase-II: Monitoring the process - Corrective and preventive action

- Univariate Control Charts
  - $\bar{X}$ - R charts or $\bar{X}$-S charts

- Multivariate Control charts
  - (i) Hotelling’s $T^2$ Control Chart - Process Mean.
    (Sub- groups, Individual Observations)
  - (ii) MCUSUM / MEWMA Control Charts - Process Variance.
Hotelling’s $T^2$ Control Chart for Individual Observations

- Sample $X_1, X_2, \ldots, X_m$ which follow a $p$-variate $MVN(\mu, \Sigma)$

\[ T^2(j) = (X_j - \bar{X})' S^{-1} (X_j - \bar{X}), j = 1, 2, \ldots, m \]  \hspace{1cm} (1)

where, $\bar{X}$ and $S$ are the sample mean & sample covariance

- $T^2$ follows $\beta$-distribution for phase-I data and $F$-distribution for phase-II data.

- Sample mean and sample covariance are sensitive to outliers.
Robust Control Charts

- Robust Control charts are suggested replacing the classical estimators by Robust estimators.
- Covariance matrix based on successive differences between vectors [Sullivan, Woodall (1996)]
- If \( V_i = X_{i+1} - X_i, i = 1, 2, \ldots, (m - 1) \), Then an unbiased estimator of covariance matrix \( S_1 = \frac{1}{2(m-1)} \sum_{i=1}^{m-1} V_i V_i' \) is replacing \( S \).
- It was effective in detecting sustained step changes but not successful in detecting multiple multivariate outliers.
Example: Comparison of Control Charts

Hoteling's T-Square Chart

T-Square Chart – Successive Difference
Robust Estimators

A Good Robust estimator should possess the following properties:

- **High Breakdown Point:**
  (i) The smallest proportion of observations which can render an estimator meaningless
  (ii) Mean has breakdown point of $1/n$ and median has $1/2$

- **Invariant:** Allows standardization/Transformation
  Highest possible asymptotic breakdown point $\frac{(m-p+1)}{2m}$

- **Efficiency:** Minimum Mean Square Error
  Trade off (Breakdown point & Efficiency)

- **Computing power in a reasonable amount of time.**
Robust Estimators

- M-estimators are robust but breakdown point reduces as dimension increases.
- Stahel-Donoho estimator with sample breakdown point is \( \frac{(m-2p+2)}{2m} \), but computationally expensive.
- S-estimator with sample breakdown point is \( \frac{(m-p+1)}{2m} \), but computationally expensive.
  (i) They are invariant
  (ii) Highest possible asymptotic breakdown point \( \frac{(m-p+1)}{2m} \).
  (iii) Fast and efficient algorithm for approximation
  (iv) Lower statistical efficiency as considers only partial data
Robust Estimators

- MVE- Find the smallest ellipsoid containing a subset \((1 - \gamma)100\%\) of the observations, \(\gamma = [0, 0.5]\)
  (i) Location - The geometrical center of the ellipsoid
  (ii) Covariance - The matrix defining the ellipsoid

- MCD- Find the subset of observations having covariance matrix with lowest determinant.
  Estimates of Mean and Covariance correspond to the mean and covariance of the subset.

- Computing exact MVE/MCD estimators are expensive but algorithms are available for approximating the estimators.
Robust Control Charts

- The $T^2$ control chart based MCD/MVE estimators [Vargas(2003), Jensen et al.(2007)]
- $T^2$ values for i-th multivariate observation is given by:

$$
T^2_{MCD}(X_i) = (X_i - \bar{X}_{MCD})' S^{-1}_{MCD} (X_i - \bar{X}_{MCD}) \\
T^2_{MVE}(X_i) = (X_i - \bar{X}_{MVE})' S^{-1}_{MVE} (X_i - \bar{X}_{MVE})
$$

- The performance of these charts are assessed by probability of signal and showed to be better than existing robust charts.
- The exact distribution of $T^2$ with MCD / MVE estimators are not available, hence the control limits are arrived by inverting empirical distribution.
- The Control limits for sample sizes ($T^2_{MVE}$ 20-75, $T^2_{MCD}$ 20-100) and dimension 2-10 at $\alpha = 0.05$ level are given.


Thank You