

3.1. CONFERENCIA PARALELA 1

Optimización Multi-objetivo, aplicando Análisis de Componentes Principales

Leonel Delgado Eraso, lederudenar@hotmail.com, Universidad de Nariño.

Resumen. El siguiente escrito contiene la definición, objetivos y propiedades de la técnica de reducción de variables mediante el análisis de componentes principales. Se define el MGI (Multivariante Global index) como una combinación lineal de los componentes principales obteniendo una función para optimizarla mediante alguna técnica mono-objetivo. Se aplicará otras técnicas como la distancia generalizada, la función de deseabilidad y el análisis envolvente de datos (DEA) a un diseño central compuesto rotatable 2^k para aplicar el GPE (Global Porcentaje Error) y comparar con la técnica del análisis de componentes principales.

Palabras claves. Optimización multi-objetivo, Optimización multi-respuesta, componentes principales, Análisis multi-respuesta.

1. Presentación.

Para la optimización mono-objetivo se han desarrollado varios métodos entre ellos: La Superficie de Respuesta, el método simplex de Nelder.Mead, el método de Programación cuadrática, el método de las direcciones factibles. El estudio de la optimización multi-objetivo tiene gran aplicación en todas las áreas del conocimiento, en la industria, en el comercio, en la ingeniería y en todas las investigaciones donde se tenga k factores y m respuestas. Varios métodos existen para la optimización multi-objetivo, la función de densidad, la función de Distancia Generalizada, el análisis envolvente de datos y métodos meta heurísticos como el algoritmo genético y las redes neuronales. Cuando el número de respuestas es muy grande los anteriores métodos se vuelven inmanejables y se recomienda usar el análisis de componentes principales como técnica de reducción de las respuestas a unas pocas que explican la mayor variabilidad, logrando con ello otras nuevas variables llamadas componentes que tienen como principal característica ser ortogonales y no están correlacionadas. Los componentes principales aplican como herramienta fundamental que son los valores propios y los vectores propios de la matriz de correlación de las respuestas, y mediante una combinación lineal de los valores propios y los componentes principales seleccionados se construye el Índice Global Multivariante (MGI, por sus sigla en inglés) para luego aplicar los métodos mono-objetivos y determinar el punto óptimo.

Se aplicaron las otras técnicas multivariantes y mediante el GPE se comparó con los resultados de componentes principales.

2. Desarrollo de la temática.

OBJETIVOS DEL PCA.

Los objetivos más importantes de todo análisis por componentes principales son:

- ✓ Generar nuevas variables que puedan expresar la información contenida en el conjunto original de datos.
- ✓ Reducir la dimensionalidad el problema que se está tratando.
- ✓ Eliminar, cuando sea posible, algunas de las variables originales ya sea porque ellas aportan poca información o porque una variable contiene en parte información ya suministrada por otra variable.

DEFINICION DEL PCA

Las nuevas variables generadas se denominan componentes principales que son combinaciones lineales de las variables originales y no están correlacionadas entre ellas. El estudio se centra en las componentes que sintetizan la mayor variabilidad del sistema de puntos.

PROPIEDADES DE LOS COMPONENTES

- 1.) Conservan la misma variabilidad inicial: La suma de las varianzas de los componentes es igual a la suma de las varianzas de las variables originales

$$\text{Var}(z_j) = \lambda_j$$

Sabemos que la suma de las varianzas en S es la suma de la traza de la matriz S , entonces

$$\text{traza}(S) = \text{var}(x_1) + \text{var}(x_2) + \dots + \text{var}(x_p) = \lambda_1 + \lambda_2 + \dots + \lambda_p$$

- 2.) La proporción de variabilidad explicada por un componente es el cociente entre su varianza (valor propio asociado al vector propio que lo define), y la suma de los valores propios de la matriz.

La proporción de la variabilidad total explicada por el componente j es:

$$\frac{\lambda_j}{\sum \lambda_j}$$

Considerando los valores propios de la matriz de correlación como un conjunto de pesos de los puntajes de PC más representativos, Paiva [12] estableció un índice global multivariado (MGI), obtenido con la suma de los productos de componentes significativos ponderados por sus respectivos valores propios. Al crear y modelar la superficie de respuesta MGI, se puede aplicar una estrategia de programación no lineal restringida, como se describe en la Ec. (12):

$$\begin{aligned} \text{Maximizar } MGI &= \sum_{i=1}^l [\lambda_i(PC_{si})] \\ \text{Sujeto a: } x^T x &\leq \rho^2 \end{aligned} \quad (12)$$

donde l es el número de componentes principales seleccionados, λ_i es el i -ésimo autovalor más grande y PC_{si} es el i th mayor PC_{score} .

Para comparar la adecuación de los modelos, se calculó el error porcentual global (GPE) para los resultados de cada método, utilizando la ecuación

$$GPE = \frac{1}{m} \sum_{i=1}^m \left| \frac{y_i^*}{\theta_i} - 1 \right|$$

donde: y_i^* son valores de las respuestas óptimas, θ_i , objetivos definidos, m , número de respuestas..

Ejemplo: Se trabajó el ejercicio 7.5 propuesto en el libro de Khuri-Cornell y se obtuvo los siguientes resultados:

	DGF	DESEABILIDAD	DEA	PCA
GPE	0,02572	0,025499	0,05527	0,027566
X_1	-0,4446	-0,4822	-0,02308	-0,3514
X_2	1,4434	1,3560	0,0783	0,8895
X_3	0,7399	0,8701	0,1658	0,6194

3. Conclusiones.

Aplicado Análisis de Componentes Principales los problemas de optimización multi-objetivo bajan la dimensionalidad en las variables respuestas.

Las aplicaciones que se han desarrollado se hicieron con diseños central compuesto rotatable 2^k pero es aplicado a cualquier otro diseño y a cualquier campos de las ciencias, donde se tenga k factores y m respuestas.

El análisis de componentes principales comparado mediante el GPE, y la distancia generalizada, la función de deseabilidad y el Análisis envolvente de datos se obtienen porcentajes similares.

4. Referencias bibliográficas.

[1] Derringer, G. y R. Suich. (3) "Simultaneous Optimization of Several Response Variables", Journal of Quality Technology, vol. 12, pp. 214-219]

[2] Montgomery, Douglas. "Diseño y análisis de experimentos". Editorial Limusa Wiley. 2004

[3] Khuri, A. I. y J. A. Cornell. Response Surfaces: Designed and Analyses. 2ª. Edición, Dekker, Nueva York.

[4] Gutierrez, Humberto. Y De La Vara Salzar, Román. "Análisis y Diseño de Experimentos" Segunda Edición.

[5] De la Vara, R. y Domínguez, J. (1998). "Metología de Superficie de Respuesta". Comunicación Técnica, I-90-08, CIMAT.

[6] Del Castillo, E.; Montgomery, D. C. and McCarville, D. R. (1996). "Modified Desirability Functions for Multiple Response Optimization". Journal of Quality Technology.

[7] Bolch, B.W., Huang, C.T., 1974. Métodos Estadísticos Multivariados para Negocios y Economía. Prentice-Hall, Acantilados de Englewood, Nueva Jersey.

[8] Charnes, A., Cooper, W.W., Rhodes, E., 1978. Medición de la eficiencia de las unidades de toma de decisiones. Revista Europea de Investigación Operativa 2, 429-444.

[9] Charnes, A. Cooper, W.W., Li, S., 1989. Utilizando el análisis de envoltura de datos para evaluar la eficiencia en el desempeño económico de las ciudades chinas. Planificación Socio-Económica 23, 325-344.