

Nenadić, O., Adler, D., Zucchini, W. (2003). Visualizing Three-Dimensional Maps in Correspondence Analysis. *CARME 2003: International Conference on Correspondence Analysis and Related Methods*. Barcelona.[Online]. Tersedia:http://www.statock.wiso.uni-oetiggen.de/mitarbeiter/ogi/pub/CARME_03_nenadic.pdf [16Juli2005]

Aam Sudrajat, Lukman, Dewi Rachmatin
Jurusan Pendidikan Matematika
FPMIPA Universitas Pendidikan Indonesia
Jl. Dr. Setiabudhi 229, Bandung 40154
E-mail address: azoery011388@yahoo.com

ABSTRAK

Analisis korespondensi menghasilkan subruang optimum dua dimensi, dari suatu baris dan kolom matriks kontingensi data berukuran $I \times J$, biasanya $I \geq J$, yang intinya mereduksi I titik pada dimensi J , sedemikian hingga menjadi I titik yang berkorespondensi dengan J titik pada subruang optimum tersebut. Kualitas informasi dari analisis korespondensi dapat ditingkatkan dengan mengambil subruang optimum tiga dimensi, yang direpresentasikan melalui bahasa pemrograman statistika **R**. Kestabilan internal merupakan kondisi di mana subruang optimum yang dihasilkan oleh analisis korespondensi tidak dipengaruhi secara signifikan oleh titik profil pencilan. Pendeteksian titik profil pencilan pada subruang tiga dimensi dapat dilakukan dengan pengkonstruksian elipsoid di mana koordinat-koordinat titik profil diasumsikan berdistribusi normal multivariat. Pencilan diperlakukan sebagai titik profil tambahan dengan pendekatan metoda Jackknifing untuk pemeriksaan kestabilan internal, dengan acuan kriteria kestabilan subruang optimum menurut Escofier dan Leroux (1976) dalam Greenacre (1984).

Kata kunci: analisis korespondensi, subruang optimum, bahasa R, elipsoid, pencilan, titik profil tambahan, Jackknifing, dan kriteria Escofier dan Le Roux (1976) dalam Greenacre (1984)

1. Pendahuluan

1.1.1 Latar Belakang Masalah

Analisis korespondensi menghasilkan subruang optimum dua atau tiga dimensi dari suatu baris dan kolom matriks data kotingensi berukuran $I \times J$, biasanya $I \geq J$, yang intinya mereduksi I titik pada dimensi J , sedemikian hingga menjadi I titik yang berkorespondensi dengan J titik pada subruang optimum. Dengan kata lain apabila digunakan subruang optimum dua dimensi maka grafik dua dimensi tersebut juga merupakan gambaran dari pencaran multidimensi titik-titik data, yang terorientasi secara optimal mencerminkan sebanyak mungkin inersia titik-titik.

Hasil analisis korespondensi dikatakan stabil secara internal jika orientasi subruang optimum tidak dipengaruhi secara signifikan oleh data pencilan. Dengan kata lain suatu sampel yang digunakan harus bebas dari suatu pengaruh pencilan, baik yang merupakan titik pencilan ataupun awan pencilan.

Pencilan pada metode penyekalaan multidimensi, khususnya analisis korespondensi, dikenal dengan sebutan titik-titik tambahan (*supplementary points*). Titik-titik tambahan dalam representasi subruang optimum yang dihasilkan dapat berasal dari kolom tambahan atau baris tambahan yang terdeteksi.

Pendeteksian pencilan yang diungkapkan oleh Greenacre (1984) adalah dengan mengeksplorasi hasil perhitungan dekomposisi inersia titik profil baris dan titik profil kolom, di mana inersia relatif titik profil yang berbeda sangat besar dari titik profil lainnya dikategorikan sebagai titik tambahan.

Masalah akan timbul apabila terdapat beberapa titik profil yang memiliki inersia relatif, sehingga tidak dapat dilakukan dengan teknik eksplorasi. Motivasi solusi yang diterapkan penulis adalah pengasumsian kenormalan distribusi dari koordinat titik-titik profil seperti yang dilakukan oleh perancang *software* STATISTICA 6.0 untuk membangun elips interval kepercayaan.

Untuk menambah kualitas informasi dari hasil analisis korespondensi maka akan dilakukan representasi subruang optimum tiga dimensi dengan pemrograman statistik berbahasa R. Kemudian seluruh analisis korespondensi akan didasarkan pada subruang optimum tiga dimensi ini.

Titik-titik profil yang terdeteksi sebagai pencilan akan diuji pengaruhnya terhadap subruang optimum dengan pendekatan metoda **Jackknifing**, yaitu dengan mengimplementasikan kriteria Escofier dan Le Roux (1976) dalam Greenacre (1984) mengenai kestabilan subruang optimum yang berotasi.

1.2 Perumusan Masalah

Berdasarkan pemaparan di atas mengenai kestabilan internal dan pengujiannya, permasalahan yang akan dibahas oleh penulis adalah:

1. Pendeteksian data pencilan yang berupa titik profil baris atau kolom dari suatu hasil analisis korespondensi dengan menggunakan elipsoid kepercayaan tiga dimensi pada taraf kepercayaan 95%.
2. Pemeriksaan kestabilan internal subruang optimum yang dihasilkan oleh analisis korepondensi dilakukan dengan mengimplementasikan kriteria Escofier dan Le Roux (1976) dalam Greenacre (1984) mengenai kestabilan subruang optimum yang berotasi ketika pencilan dijadikan titik profil tambahan.
3. Perbandingan hasil pengujian kestabilan subruang optimum tiga dimesi dengan pemeriksaan kestabilan subruang optimum dua dimensi yang diperoleh Greenacre (1984).

2. Isi

2.1 Pendeteksian Pencilan dengan Menggunakan Elipsoid Kepercayaan

Jika $\mathbf{X} \sim N_p(\mathbf{0}, \Sigma)$ dengan Σ matriks simetris dan definit positif maka $(\mathbf{X} - \boldsymbol{\mu})' \Sigma^{-1} (\mathbf{X} - \boldsymbol{\mu}) \sim \chi_p^2$, di mana kontur dari kepadatan peluang konstan untuk distribusi normal p -variat adalah elipsoid yang didefinisikan oleh \mathbf{X} sedemikian hingga $(\mathbf{X} - \boldsymbol{\mu})' \Sigma^{-1} (\mathbf{X} - \boldsymbol{\mu}) = c^2$. Elipsoid kepercayaan yang memiliki taraf kepercayaan $(1 - \alpha)100\%$, yaitu $P\left[(\mathbf{X} - \boldsymbol{\mu})' \Sigma^{-1} (\mathbf{X} - \boldsymbol{\mu}) \leq \chi_p^2(\alpha)\right] = 1 - \alpha$.

Penerapan elipsoid kepercayaan pada sampel, diharapkan $(1 - \alpha)\%$ data dari sampel akan terletak pada elipsoid $(\mathbf{X} - \bar{\mathbf{X}})' \mathbf{S}^{-1} (\mathbf{X} - \bar{\mathbf{X}}) \leq \chi_p^2(\alpha)$, di mana $\bar{\mathbf{X}}$ dan \mathbf{S} merupakan penaksir-penaksir dari $\boldsymbol{\mu}$ dari \mathbf{S} yang diperoleh dari sampel.

Pandang koordinat-koordinat titik profil baris dan kolom yang dihasilkan untuk subruang tiga dimensi, yaitu:

$$\mathbf{F}_{(3)} = \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ \vdots & \vdots & \vdots \\ f_{J1} & f_{J2} & f_{J3} \end{bmatrix} \text{ dan } \mathbf{G}_{(3)} = \begin{bmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ \vdots & \vdots & \vdots \\ g_{J1} & g_{J2} & g_{J3} \end{bmatrix}.$$

Dengan asumsi kenormalan distribusi dari $\mathbf{F}_{(3)}$ dan $\mathbf{G}_{(3)}$, yaitu distribusi normal multivariat (trivariat) maka elipsoid kepercayaan untuk sampel $P\left[(\mathbf{X} - \bar{\mathbf{X}})' \mathbf{S}^{-1} (\mathbf{X} - \bar{\mathbf{X}}) \leq \chi_p^2(\alpha)\right] = 1 - \alpha$ dapat diimplementasikan pada $\mathbf{F}_{(3)}$ dan $\mathbf{G}_{(3)}$, sedemikian hingga diperoleh

$$P\left[(\mathbf{F}_{(3)} - \bar{\mathbf{F}}_{(3)})' \mathbf{S}_{\mathbf{F}_{(3)}}^{-1} (\mathbf{F}_{(3)} - \bar{\mathbf{F}}_{(3)}) \leq \chi_3^2(\alpha)\right] = 1 - \alpha \text{ dan}$$

$$P\left[(\mathbf{G}_{(3)} - \bar{\mathbf{G}}_{(3)})' \mathbf{S}_{\mathbf{G}_{(3)}}^{-1} (\mathbf{G}_{(3)} - \bar{\mathbf{G}}_{(3)}) \leq \chi_3^2(\alpha)\right] = 1 - \alpha$$

yang masing-masing menunjukkan elipsoid kepercayaan $(1 - \alpha)100\%$ untuk titik-titik profil baris dan kolom dalam subruang optimal tiga dimensi di mana titik-titik profil tersebut direpresentasikan. Dengan demikian titik profil yang di luar dari elipsoid kepercayaan merupakan titik profil pencilan.

2.2 Pendekatan Jackknifing

Pendekatan Jackknifing untuk menguji kestabilan subruang dilakukan dengan menghilangkan satu titik profil baris atau kolom, kemudian dilakukan perhitungan ulang dari orientasi bidang yang dihasilkan melalui perhitungan koordinat-koordinat baris dan kolom serta inersianya.

Bila dimiliki suatu matriks kontingensi berukuran $(I \times J)$, pendekatan Jackknifing yang dilakukan akan banyak menyita waktu dengan melakukan analisis sebanyak I kali untuk menghitung orientasi subruang optimal setiap satu titik profil baris dihilangkan. Untuk meringkas pendekatan Jackknifing ini maka digunakan elipsoid kepercayaan sebagai pengganti metoda eksplorasi hasil dekomposisi inersia seperti yang dilakukan Greenacre (1984). Titik profil yang berada di luar elipsoidlah yang merupakan pencilan. Pencilan tersebut diuji terhadap bidang dengan menghilangkannya dari analisis, kemudian dilakukan perhitungan ulang koordinat-koordinat baris dan kolom serta dekomposisi inersianya.

2.3 Kriteria Kestabilan Subruang Optimum Dua Dimensi Menurut Escofier dan Le Roux (1976) dalam Greenacre (1984)

Misalkan suatu subruang optimum yang dipengaruhi oleh suatu pencilan (calon titik profil tambahan), kemudian pencilan tersebut dijadikan titik profil tambahan, jika terjadi pertukaran sumbu utama ke- k dan ke- $k+1$ maka subruang optimum tersebut tidak stabil.

Jika sebuah titik profil baris $i = s$ dihilangkan maka subruang optimum akan memiliki sebuah sentroid baru dan pencocokan terbaik (*best-fitting*) bidang akan memiliki orientasi yang baru ketika bidang tersebut tidak dipengaruhi oleh posisi s . Tergantung daripada ukuran kontribusi-kontribusi r_{s,s_1}^2 dan r_{s,s_2}^2 , salah satu dari situasi berikut dapat dihasilkan:

- (1) Bidang tersebut dan sumbu-sumbu menjadi stabil
- (2) Bidang tersebut menjadi stabil dan sumbu 1 dan sumbu 2 berubah orientasi, bahkan saling tukar.
- (3) Bidang tersebut menjadi stabil sebagian, sebagai contoh sumbu 1 stabil tetapi sumbu 2 memutar bidang, bahkan bertukar dengan sumbu 3.
- (4) Bidang tersebut tidak stabil dan diasumsikan sebuah orientasi baru yang sempurna.

Menurut Escofier dan Le Roux (1976) dalam Greenacre (1984) bila sudut rotasi bidang dua dimensi adalah ϕ maka batas atas rotasi bidang dua dimensi dapat dinyatakan dalam bentuk dalam bentuk :

$$\sin 2\phi \leq \lambda_k \left[\frac{\alpha - \beta}{\varepsilon(1 + \beta)} \right]$$

di mana, dalam penotasian biasa dari analisis korespondensi dan pengasumsian bahwa baris ke- s dari matriks proporsi P dihilangkan:

Jika yang diselidiki adalah kestabilan ruang dua dimensi hasil analisis korespondensi maka $\varepsilon_2 \equiv \min\{\lambda_1 - \lambda_2, \lambda_2 - \lambda_3\}$ dan jika nilai λ_{k+1} dekat dengan λ_k sedemikian hingga maka Escofier dan Le Roux (1976) menyarankan menggunakan $\varepsilon \equiv \min\left\{\lambda_{k-1} - \lambda_k, \frac{1}{2}(\lambda_{k+1} - \lambda_{k+2})\right\}$.

Subruang optimum 2 dimensi dikatakan stabil secara internal bila nilai ϕ tidak melebihi 45° .

2.4 Visualisasi 3 Dimensi Analisis Korespondensi dengan Bahasa Pemrograman R

Visualisasi tiga dimensi pada analisis korespondensi menambahkan informasi mengenai massa dari suatu titik profil baris atau kolom, sehingga dengan menampilkan massa pada plot tiga dimensi tersebut dapat diketahui penentu orientasi subruang optimal tiga dimensi tersebut dan dapat diinterpretasikan bahwa titik profil dengan massa yang besar berkorespondensi sangat kuat dengan titik profil lain didekatnya.

Visualisasi tiga dimensi dengan menggunakan bahasa R membutuhkan paket RGL yang dikembangkan oleh para pengembang bahasa R. *RGL device* memberikan suatu ruang tiga dimensi seperti halnya program pengolah grafik tiga dimensi 3Dmax.

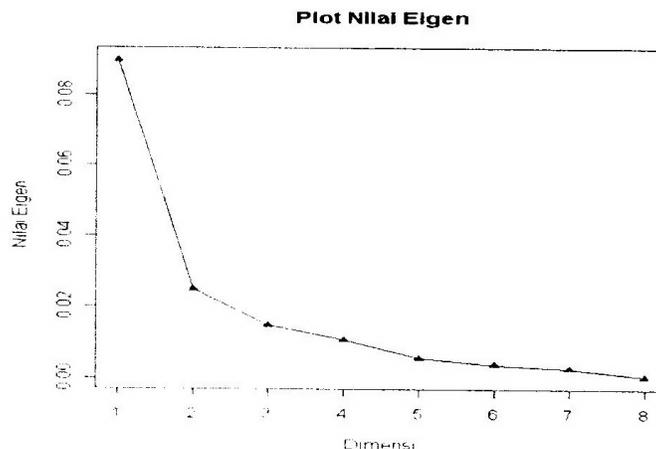
Representasi titik-titik profil baris dan kolom menggunakan paket RGL memungkinkan untuk dilakukan pandangan menyeluruh terhadap ruang tiga dimensi yang dihasilkan dengan melakukan perbesaran dan rotasi secara langsung dengan *mouse* atau fungsi *rgl.viewpoint()* yang ada. Dengan melakukan perbesaran dan rotasi pandangan, eksplorasi dan penyelidikan korespondensi antara titik profil baris dan kolom dapat dilakukan secara menyeluruh.

Pada tahap implementasi program digunakan selimut bola semi transparan berjari-jari r yang merepresentasikan massa dari titik profil, di mana jari-jari (r) bola merupakan fungsi dari massa baris dan kolom yang bersesuaian untuk setiap titik profil.

2.5 Studi Kasus

Dengan menggunakan data tabulasi silang antara konsumsi protein per orang per gram per hari dengan negara-negara di Eropa pada tahun 1973 dari Weeber (1973) dalam Greenacre (1984) dan fungsi-fungsi yang dibangun dengan bahasa **R** untuk merepresentasikan korespondensi dan mendeteksi titik profil pencilan diperoleh plot tiga dimensi hasil analisis korespondensi seperti pada gambar 1, di mana persamaan elipsoidnya adalah

$$\left(\frac{f_{1i} + 0.00147}{0.8423273}\right)^2 + \left(\frac{f_{2i} + 0.0046}{0.5833602}\right)^2 + \left(\frac{f_{3i} + 0.0263}{0.410542}\right)^2 = 1.$$



Gambar 1. Plot Nilai Eigen di mana Titik Profil PRT Menjadi Titik Profil Baris Tambahan

Pada gambar 3 di atas, titik-titik berwarna hijau merupakan titik-titik profil baris sebelum titik profil baris PRT dijadikan titik profil tambahan baris, sedangkan titik-titik berwarna biru adalah titik-titik profil baris sesudah titik profil baris PRT dijadikan titik profil baris tambahan. Dari gambar tersebut terlihat bahwa titik-titik profil baris sebelum dan sesudah titik profil baris PRT dijadikan titik profil tambahan hanya menunjukkan sedikit perubahan, dengan kata lain kondisi kumpulan (awan) titik profil baris relatif tetap. Untuk melihat kondisi kumpulan titik profil kolom maka dilakukan hal yang sama seperti pada titik-profil baris, dan diperoleh hasil yang identik.

Dari seluruh pembahasan diatas dapat diinterpretasikan bahwa subruang optimal tiga dimensi yang dihasilkan dari analisis korespondensi terhadap data tabulasi silang antara konsumsi protein per orang per gram per hari dengan negara-negara di Eropa pada tahun 1973 dari Weeber (1973) dalam Greenacre (1984) adalah **stabil**, dengan kata lain subruang optimal tiga dimensi yang dihasilkan memiliki kestabilan internal yang cukup berarti.

3. Kesimpulan dan Saran

3.1 Kesimpulan

Dari tugas akhir penulisan mengenai “KESTABILAN INTERNAL DALAM ANALISIS KORESPONDENSI” ini, berikut ini adalah hal-hal yang dapat disimpulkan:

1. Pendeteksian titik profil pencilan menggunakan elipsoid kepercayaan dengan taraf kepercayaan 95% dengan asumsi kenormalan distribusi dari koordinat titik-titik profil, dapat digunakan sebagai alternatif teknik eksplorasi hasil dekomposisi inersia.
2. Pemeriksaan subruang tiga dimensi dilakukan dengan eksplorasi ruang tiga dimensi menggunakan bantuan bahasa pemrograman statistik **R**, di mana hasil sebelum dan sesudah titik profil pencilan dijadikan titik profil tambahan dapat diplot secara bersamaan dan diamati perubahan konsidi kumpulan titik profilnya.
3. Perbandingan hasil pemeriksaan kestabilan subruang tiga dimensi dengan hasil Greenacre(1984) pada subruang optimum dua dimensi untuk data konsumsi protein 25 negara dan 9 sumber protein seperti tercantum pada tabel 4.1 menunjukkan hasil yang cukup identik.

3.2 Saran

1. Pengujian pencilan dengan elipsoid kepercayaan hendaknya dilakukan pada data yang diperkirakan berdistribusi normal dan berukuran cukup besar, untuk meyakinkan dapat dilakukan uji pendahuluan dari kenormalan distribusi koordinat-koordinat titik profil.
2. Pengujian kestabilan pada ruang tiga dimensi yang dilakukan penulis belum dapat divisualisasikan, sehingga untuk kemajuan analisis korespondensi sebagai suatu teknik statistika deskriptif, hendaknya ada penelitian lebih lanjut.
3. Penggunaan *software* bahasa pemrograman statistik **R**, khususnya dengan menggunakan paket *library* RGL dan misc3D, hendaknya dilakukan pada kartu grafis yang mendukung 3D dan didukung oleh RAM dan processor yang handal. Sebagai contoh penulis menggunakan processor 2.26 GHz, 256 DDR RAM, 64 MB VGA card.

Daftar Pustaka

- Greenacre, M.J. (1984). *Theory And Application Correspondence Analysis*. London: Academic Press.
- Lukman, dan Marwati, R. (2004). *Metoda Analisis Korespondensi Dalam Mengolah Data Kuisisioner*. Bandung: Jurusan Pendidikan Matematika UPI.
- Nenadic, O. (onenadi@uni-goettingen.de). (2005,14 Juli). AW: *Ellipse Confidence/Bound in Correspondence Analysis*. E-mail kepada Aam Sudrajat(azoery011388@yahoo.com)



PROSIDING SEMINAR NASIONAL MATEMATIKA

Peranan Matematika dalam Pengembangan
Teknologi Informasi dan Komunikasi

Bandung, 20 Agustus 2005



Diselenggarakan Oleh:

**JURUSAN PENDIDIKAN MATEMATIKA
FAKULTAS PENDIDIKAN MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS PENDIDIKAN INDONESIA
2005**