UNIVERSITAS INDONESIA

PEMBENTUKAN INDIKATOR SASARAN DENGAN PROXY MEANS TEST BERDASARKAN METODE PRINCALS

SKRIPSI

WINDA JUWITA SARI
0706163161

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
PROGRAM STUDI SARJANA MATEMATIKA
DEPOK
JULI 2011

UNIVERSITAS INDONESIA

PEMBENTUKAN INDIKATOR SASARAN DENGAN PROXY MEANS TEST BERDASARKAN METODE PRINCALS

SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana sains

WINDA JUWITA SARI
0706163161

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
PROGRAM STUDI SARJANA MATEMATIKA
DEPOK
JULI 2011

HALAMAN PERNYATAAN ORISINALITAS

Skripsi ini adalah hasil karya sendiri,

dan semua sumber baik yang dikutip maupun dirujuk

telah saya nyatakan dengan benar.

Nama	: Winda Juwita Sari
NPM	: 0706163161
Tanda Tangan	:
Tanggal	: 8 Juli 2011

HALAMAN PENGESAHAN

Skripsi ini diajukan oleh

Nama	: Winda Juwita Sari
NPM	: 0706163161
Program Studi	: Sarjana Matematika
Judul Skripsi	: Pembentukan Indikator Sasaran dengan *Proxy Means Test* Berdasarkan Metode PRINCALS

Telah berhasil dipertahankan di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Sains pada Program Studi Sarjana Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Indonesia

DEWAN PENGUJI

Pembimbing	: Dra. Rianti Setiadi, M.Si	( )
Penguji	: Dr. Dian Lestari	( )
Penguji	: Fevi Novkaniza, S.Si, M.Si	( )
Penguji	: Dra. Rustina	( )

Ditetapkan di	: Depok
Tanggal	: 10 Juni 2011

KATA PENGANTAR

Puji syukur kepada Allah SWT atas segala berkah dan karunia-Nya sehingga penulisan skripsi ini dapat diselesaikan dengan baik. Penulisan skripsi ini dilakukan dalam rangka memenuhi salah satu syarat untuk mencapai gelar Sarjana Sains Departemen Matematika pada Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Indonesia.

Penulis menyadari bahwa penyelesaian penulisan skripsi tidak terlepas dari bantuan dan dukungan dari berbagai pihak. Oleh karena itu, penulis menghaturkan terima kasih kepada:

(1) Dra. Rianti Setiadi, M.Si selaku pembimbing skripsi yang telah banyak meluangkan waktu untuk membimbing, memberi saran, dan memberi bantuan untuk penulis dalam menyelesaikan tugas akhir ini.

(2) Dra. Ida Fithriani M.Si selaku pembimbing akademis penulis yang telah memberi pengarahan dan dukungan selama menjalani masa kuliah.

(3) Ayah dan ibuku tercinta atas semua kasih sayang, doa, dukungan, dan kepercayaan yang telah engkau berikan selama ini. Kalianlah orang yang paling sempurna di mataku dan juga sebagai inspirator dalam kehidupanku.

(4) Kakakku tercinta, Bang Zainal dan Bang Chairul atas semua dukungan, kepercayaan, dan perlindungan yang telah kuterima.

(5) Mbak Eka, keponakanku tercinta Galang, Pak Ethe Iyus, Pak Ethe Sichan, Ethe Emi, dan seluruh keluarga besar penulis yang telah memberikan doa, dukungan, baik moril maupun materi.

(6) Bapak dan Ibu dosen yang telah hadir dan memberikan saran-saran kepada penulis mulai dari sig 1 sampai kolokium dan sidang, Ibu Saskya Mary, Ibu Siti Nurrohmah, Ibu Rustina, Mba Sarini Abdullah, Mba Mila Novita, Bu Dian, Mba Fevi dan Bu Yekti.

(7) Seluruh dosen beserta staf Departemen Matematika FMIPA UI atas bantuan dan bimbingannya.

(8) Kaka' atas kebersamaan, dukungan, doa, dan motivasi. Maaf ya sering menyusahkan dan kekanak-kanakan.

(9) Farah, Widi, Lois, Nora, Dita, Toto, Bapet, Kiki, dan Ferdi atas kebersamaan dan dukungan kalian selama ini. Kebersamaan kita mulai dari makan bareng, nonton bareng, jalan-jalan, sampe liburan bareng akan jadi kenangan indah. Jangan lupa rencana kita jenguk Lois di Perancis. hehehe

(10) Teman-teman angkatan 2007, Nedi, Shafira, Widya, Hikmah, Manda, Adit, Danar, Andi, Hanif, Ayat, Adi, Param, Sisca, Shafa, Gamar, Anjar, Siska, Putu, Isna, Bowo, Anggun, Zul, Anis, Yos, Siti, Misda, Afni, Fauzan, Tepi, dan Wiwi atas dukungan dan kebersamaan mulai dari awal kuliah sampai akhir kuliah.

(11) Ka Ajat yang telah membantu penulis saat kesulitan menurunkan rumus. Ka Amri dan Ka May yang telah mempercayakan penulis menjadi asisten lab statistik.

(12) Kakak-kakak angkatan 2006, Ka Teguh, Ka Poe, Ka Ali, Ka Rendy, Ka Billy, Ka Farah, Ka Reza, Ka Bekti, Ka Nita, Ka Rizky, Ka AR, Ka Indah, dan lain-lain atas bimbingannya selama ini.

(13) Angkatan 2008, Luthfa, Dheni, Numa, Andi, Adi, Olin, Cindy, Emy, Eka, dan lain-lain. Selamat berjuang dan semoga cepat lulus ya

(14) Adik asuhku, Dinda dan Icha, dan angkatan 2009, Ozan, Alfian, Azki, Revi, dan lain-lain. Semangat kuliahnya

(15) Seluruh teman-teman angkatan 2010

Penulis juga mengucapkan terima kasih kepada pihak yang tidak dapat disebutkan satu per satu, yang telah membantu dalam penyusunan tugas akhir ini. Akhir kata, penulis mohon maaf apabila terdapat kesalahan atau kekurangan dalam tugas akhir ini. Semoga tugas akhir ini membawa manfaat bagi perkembangan ilmu pengetahuan.

Penulis

2011

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI
TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Universitas Indonesia, saya yang bertanda tangan di bawah ini:

Nama	: Winda Juwita Sari
NPM	: 0706163161
Program Studi	: Sarjana Matematika
Departemen	: Matematika
Fakultas	: Matematika dan Ilmu Pengetahuan Alam
Jenis karya	: Skripsi

demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Indonesia Hak Bebas Royalti Noneksklusif (*Non-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul :

Pembentukan Indikator Sasaran dengan *Proxy Means Test* Berdasarkan Metode PRINCALS.

beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Noneksklusif ini Universitas Indonesia berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Depok
Pada tanggal : 8 Juli 2011
Yang menyatakan

(Winda Juwita Sari)

ABSTRAK

Nama : Winda Juwita Sari
Program Studi : S1 Matematika
Judul : Pembentukan Indikator Sasaran dengan *Proxy Means Test* Berdasarkan Metode PRICALS

Indikator sasaran digunakan untuk mengukur suatu variabel sasaran yang tidak bisa diukur secara langsung dan tidak bisa diukur dengan skala likert. Salah satu cara membentuk indikator sasaran adalah *proxy means test*, yaitu membentuk indikator sasaran dari sejumlah variabel proxi yang berhubungan erat dengan variabel sasaran, mudah diukur, dan tidak bisa dimanipulasi. Salah satu metode yang digunakan dalam *proxy means test*, di mana variabel-variabel proxi merupakan variabel kategorik, adalah metode PRINCALS. Prinsip dari PRINCALS adalah kuantifikasi dengan teknik *optimal scaling* dan pembentukan sejumlah komponen dari variabel-variabel kategorik yang telah dikuantifikasi, di mana dilakukan secara iteratif dan simultan sehingga memaksimumkan total variansi variabel-variabel kategorik yang telah dikuantifikasi yang dicakup oleh sejumlah komponen utama. Indikator sasaran *proxy means test* berdasarkan metode PRINCALS dipilih dari beberapa komponen yang telah terbentuk sesuai kebutuhan peneliti.

Kata kunci : indikator sasaran, *proxy means test*, *optimal scaling*, PRINCALS
xiii+60 halaman : 17 tabel
Daftar Pustaka : 15 (1978-2010)

ABSTRACT

Name : Winda Juwita Sari
Study Program: S1 Mathematics
Title : Establisment of Target Indicator by Proxy Means Test Based on PRINCALS Method

Target indicators are used to measure a target variable that can not be measured directly and can not be measured with a Likert scale. One way to form a target indicators are with proxy means test, which form the target indicator of the number of proxy variables which are closely related to the target variable, easily measured, and can not be manipulated. One of the methods used in proxy means test, in which the proxy variables are categorical variables, is the PRINCALS method. The principle of PRINCALS is the quantification with optimal scaling technic and establisment of components of categorical variables that have been quantified, which is done iteratively and simultaneously to maximize the total variance of categorical variables covered by a number of major components. Target indicators of proxy means test based on PRINCALS method are chosen from several components that have been formed according to the needs of researchers.

Key words : indicator of targeting, proxy means test, optimal scaling, PRINCALS
xiii+60 pages : 17 tables
Bibliography : 15 (1978-2010)

DAFTAR ISI

DAFTAR TABEL

DAFTAR LAMPIRAN

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Di dalam beberapa penelitian sosial, terutama dalam menentukan sasaran program sosial, variabel yang menjadi sasaran program sosial tersebut biasanya tidak bisa diukur secara langsung dan tidak bisa juga diukur menggunakan skala Likert. Misalkan, program Raskin untuk masyarakat miskin, di mana variabel tingkat kemiskinan merupakan variabel sasaran yang tidak bisa diukur secara langsung. Untuk mengatasi masalah tersebut, perlu ditentukan suatu indikator yang bisa diukur secara langsung dan secara subtantif mencerminkan variabel sasaran. Selanjutnya, indikator tersebut akan disebut sebagai indikator sasaran. Namun, indikator sasaran tersebut kadang kala tidak bisa dipercaya (bisa dimanipulasi). Sebagai contoh, misalnya indikator yang dipilih untuk mencerminkan tingkat kemiskinan adalah pendapatan keluarga per bulan, tetapi saat pengambilan sampel, responden memanipulasi pendapatan mereka karena mengetahui tujuan dari pengambilan sampel ini adalah menentukan sasaran penerima Raskin. Salah satu metode yang digunakan untuk mengatasi masalah tersebut adalah *Proxy Means Test*. *Proxy Means Test* adalah suatu cara untuk mencari suatu indikator sasaran yang dibangun dari beberapa variabel yang secara subtantif berhubungan erat dengan variabel sasaran. Variabel-variabel tersebut disebut sebagai variabel proxi. Variabel-variabel proxi yang dipilih adalah variabel-variabel yang mudah diukur dan tidak bisa dimanipulasi pada saat pengukuran.

Salah satu metode statistik yang biasanya digunakan dalam *Proxy Means Test* adalah metode regresi linier, dengan variabel responnya adalah indikator sasaran yang masih mungkin dimanipulasi, dan variabel prediktornya adalah variabel-variabel proxi. Kemudian, digunakan taksiran variabel respon yang didapat dari model sebagai indikator sasaran yang digunakan untuk mewakili variabel sasaran. Kelemahan dari metode ini adalah saat membentuk persamaan regresi linier masih menggunakan data dari indikator sasaran yang masih mungkin dimanipulasi. Untuk mengatasi hal ini, sering digunakan metode lain dalam *proxy means test*, yaitu dengan membangun indikator sasaran menggunakan Analisis Komponen Utama. Biasanya, indikator sasaran akan dipilih dari komponen utama yang pertama, yaitu komponen utama yang mengandung informasi maksimum dari variabel-variabel proxi. Tetapi, jika informasi variabel-variabel proxi yang dicakup dalam komponen utama yang pertama dianggap kurang memadai, maka komponen utama kedua, ketiga dan seterusnya secara bersama-sama akan dipertimbangkan untuk dipilih sebagai indikator sasaran. Dengan demikian, data dari indikator sasaran yang masih dapat dimanipulasi tidak digunakan dalam membentuk indikator sasaran yang akan digunakan untuk mewakili variabel sasaran.

Hal lain yang dihadapi dalam penelitian sosial adalah bahwa variabel-variabel proxi seringkali merupakan variabel katagorik. Hal ini menjadi suatu kendala saat menggunakan analisis komponen utama karena dalam membentuk komponen utama, variabel-variabel yang dilibatkan adalah variabel numerik. Untuk itu, diperlukan pemberian nilai numerik atau kuantifikasi yang tepat untuk setiap kategori pada setiap variabel kategorik karena pemberian nilai numerik yang berbeda akan memberikan hasil analisis yang berbeda. Salah satu metode yang bisa digunakan untuk menentukan nilai numerik pada setiap kategori untuk setiap variabel kategorik yang dilibatkan dalam analisis komponen utama adalah menggunakan prinsip *optimal scaling*, yaitu memberikan nilai numerik kepada setiap kategori untuk setiap variabel kategorik sehingga memaksimumkan total variansi variabel-variabel proxi yang telah dikuantifikasi yang dicakup oleh banyaknya komponen utama yang dipilih. Analisis komponen utama yang menggunakan prinsip *optimal scaling* disebut PRINCALS.

Pada tugas akhir ini, penulis akan membahas tentang pembentukan indikator sasaran dengan *proxy means test* berdasarkan metode PRINCALS. Selanjutnya, metode tersebut akan diterapkan untuk menentukan indikator sasaran tingkat kemiskinan berdasarkan enam variabel proxi yang telah diteliti oleh suatu lembaga yang memperhatikan penanggulangan kemiskinan di Indonesia sebagai variabel yang berhubungan erat dengan tingkat kemiskinan di empat provinsi di Indonesia, yaitu Kalimantan Barat, Nusa Tenggara Barat, Sulawesi Selatan, dan Maluku. Keenam variabel proxi tersebut adalah jenis lantai terluas, jenis dinding terluas, jenis atap, pemakaian jamban/kaskus (sendiri/ bersama), sumber penerangan utama, dan bahan bakar utama. Selanjutnya, berdasarkan indikator sasaran tingkat kemiskinan tiap provinsi, akan dilihat apakah data total penghasilan rumah tangga di keempat provinsi tersebut termanipulasi atau tidak.

1.2 Permasalahan

Bagaimanakah membentuk indikator sasaran dengan *proxy means test* berdasarkan metode PRINCALS.

1.3 Tujuan

Membentuk indikator sasaran dengan *proxy means test* berdasarkan metode PRINCALS.

1.4.1 Sistematika Penulisan

Sistematika penulisan pada tugas akhir ini adalah sebagai berikut:

Bab 1 Pendahuluan

Terdiri dari latar belakang, permasalahan, tujuan penulisan, dan sistematika penulisan.

Bab 2 Landasan Teori

Berisi uraian mengenai teori-teori yang mendukung isi tugas akhir ini yang meliputi teori mengenai Analsis Komponen Utama (Pembentukan Komponen Utama, dan Penentuan Jumlah Komponen Utama), dan teknik *optimal scaling*.

Bab 3 Pembentukkan Indikator Sasaran dengan *Proxy Means Test* Berdasarkan Metode PRINCALS

Berisi penjelasan mengenai tahapan-tahapan dalam metode PRINCALS dan cara membentuk indikator sasaran dengan *proxy means test* berdasarkan metode PRINCALS.

Bab 4 Contoh Penerapan

Berisi suatu contoh mengenai pembentukan indikator sasaran tingkat kemiskinan menggunakan enam variabel proxi, yaitu jenis lantai terluas, jenis dinding terluas, jenis atap, pemakaian jamban/kaskus (sendiri/ bersama), sumber penerangan utama, dan bahan bakar utama, untuk empat provinsi, yaitu Kalimantan Barat, Nusa Tenggara Barat, Sulawesi Selatan, dan Maluku. Selanjutnya, berdasarkan indikator sasaran tingkat kemiskinan tiap provinsi, akan dilihat apakah data total penghasilan rumah tangga di keempat provinsi tersebut termanipulasi atau tidak.

Bab 5 Penutup

Berisi kesimpulan dan saran.

BAB 2

LANDASAN TEORI

Pada bab ini akan dibahas teori-teori yang mendukung isi tugas akhir ini yang meliputi teori mengenai Analsis Komponen Utama (Pembentukan Komponen Utama, dan Penentuan Jumlah Komponen Utama), dan teknik *optimal scaling*.

2.1 Analisis Komponen Utama

Dalam melakukan penelitian, seringkali peneliti harus melibatkan atau menganalisis sejumlah variabel yang cukup banyak, di mana variabel-variabel tersebut saling berkorelasi satu dengan yang lain. Jumlah variabel yang cukup banyak, menuntut sejumlah data yang juga cukup banyak, dan peneliti kadang kala mengalami kesulitan saat menginterpretasikan hasil analisis. Selain itu, sebagian besar metode statistik yang digunakan dalam analisis, mengasumsikan tidak adanya korelasi antara variabel-variabel yang dilibatkan dalam analisis, terutama variabel-variabel yang dijadikan variabel prediktor. Oleh karena itu, diperlukan suatu metode yang bisa digunakan untuk mereduksi banyaknya variabel tersebut (selanjutnya disebut sebagai variabel asal) menjadi sejumlah variabel baru, yang tidak berkorelasi satu dengan yang lain, dan mencakup semaksimal mungkin informasi dari variabel-variabel asal.

Salah satu metode statistik yang bisa digunakan untuk mencapai tujuan tersebut adalah metode Analisis Komponen Utama (*Principal Component Analysis*). Analisis komponen utama adalah metode statistik yang digunakan untuk membentuk variabel baru yang merupakan kombinasi linier dari variabel-variabel asal, di mana variabel-variabel baru tersebut tidak berkorelasi satu dengan yang lain. Dalam analisis komponen utama, variabel baru yang terbentuk disebut sebagai komponen utama. Informasi yang diberikan oleh variabel-variabel asal diwakili oleh variansi dari variabel-variabel asal tersebut, sedangkan informasi yang diberikan oleh komponen-komponen utama diwakili oleh variansi dari komponen utama yang terbentuk. Banyaknya komponen utama maksimum yang bisa dibentuk adalah sama dengan banyaknya variabel asal. Namun, karena tujuan dari penggunaan analisis komponen utama adalah mereduksi banyaknya variabel-variabel asal menjadi sejumlah komponen utama, maka tidak semua komponen utama akan diperhitungkan dalam analisis. Banyaknya komponen utama yang akan diperhitungkan, ditentukan dengan mempertimbangkan persentasi dari total variansi dari variabel-variabel asal yang sudah tercakup dalam komponen-komponen utama.

2.1.1 Pembentukan Komponen-Komponen Utama

Misalkan, X_1, X_2, …, X_p merupakan variabel-variabel acak. Sebut, $\boldsymbol{X}^t = (X_{\boldsymbol{1}} \quad X_{\boldsymbol{2}} \quad \dots \quad X_{\boldsymbol{p}})$. Asumsikan, $E(X_i) = 0$, untuk setiap i = 1, 2, …, p, sehingga $E(\boldsymbol{X}^t) = \big(E(X_{\boldsymbol{1}}) \quad E(X_{\boldsymbol{2}}) \quad \dots \quad E(X_{\boldsymbol{p}})\big) = (0 \quad 0 \quad \dots \quad 0) = \boldsymbol{0}^t$. Sebut, $E(\boldsymbol{X}\boldsymbol{X}^t) = \boldsymbol{\Sigma}$. Matriks $\boldsymbol{\Sigma}$ merupakan matriks variansi-kovariansi dari X_1, X_2, …, X_p.

Tujuan dari analisis komponen utama adalah mencari komponen-komponen utama, sebut F_1, F_2, …, F_p, yang merupakan kombinasi linier dari X_1, X_2, …, X_p, sedemikian sehingga $var(F_1) \geq var(F_2) \geq \cdots \geq var(F_p)$, dan F_1, F_2, …, F_p tidak berkorelasi satu dengan yang lain.

Misalkan, $\lambda_1, \lambda_2, \dots, \lambda_p$ adalah nilai eigen dari matriks $\boldsymbol{\Sigma}$, di mana $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p$. Karena $\boldsymbol{\Sigma}$ marupakan matriks yang simetri berukuran (p x p), dapat dibuktikan bahwa matriks $\boldsymbol{\Sigma}$ memiliki p vektor eigen yang ortonormal (lampiran). Misalkan, $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \dots, \boldsymbol{\alpha}_p$ merupakan vektor eigen dari matriks $\boldsymbol{\Sigma}$, yang bersesuaian dengan nilai eigen $\lambda_1, \lambda_2, \dots, \lambda_p$, maka $\boldsymbol{\alpha}_i^t\boldsymbol{\alpha}_j = 0$, dan $\boldsymbol{\alpha}_i^t\boldsymbol{\alpha}_i = 1$. Vektor eigen $\boldsymbol{\alpha}_i$, dan nilai eigen λ_i memenuhi persamaan $\boldsymbol{\Sigma}\boldsymbol{\alpha}_i = \lambda_i\boldsymbol{\alpha}_i$, untuk setiap i = 1, 2, …, p.

Bentuk $F_i = \boldsymbol{\alpha}_i^t\boldsymbol{X}$, untuk setiap i = 1, 2, …, p. Akan dibuktikan bahwa $var(F_1) \geq var(F_2) \geq \cdots \geq var(F_p)$.

Bukti :

Karena mean dari F_i adalah $E(F_i) = E(\boldsymbol{\alpha}_i^t \boldsymbol{X}) = \boldsymbol{\alpha}_i^t E(\boldsymbol{X}) = 0$, untuk setiap i = 1, 2, …, p, maka variansi dari F_i dapat ditulis sebagai berikut.

$$
\begin{aligned}
var(F_i) &= E(F_i F_i^t) \\
&= E(\boldsymbol{\alpha}_i^t \boldsymbol{X} (\boldsymbol{\alpha}_i^t \boldsymbol{X})^t) \\
&= \boldsymbol{\alpha}_i^t E(\boldsymbol{X}\boldsymbol{X}^t) \boldsymbol{\alpha}_i \\
&= \boldsymbol{\alpha}_i^t \boldsymbol{\Sigma} \boldsymbol{\alpha}_i \\
&= \boldsymbol{\alpha}_i^t \lambda_i \boldsymbol{\alpha}_i \\
&= \lambda_i \boldsymbol{\alpha}_i^t \boldsymbol{\alpha}_i = \lambda_i \qquad (2.1.1)
\end{aligned}
$$

Karena $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p$, dan $var(F_i) = \lambda_i$, untuk setiap i = 1, 2, …, p, maka terbukti bahwa $var(F_1) \geq var(F_2) \geq \cdots \geq var(F_p)$.

Selanjutnya, akan dibuktikan bahwa F_1, F_2, …, F_p tidak berkorelasi satu dengan yang lain.

Bukti :

$$
\begin{aligned}
cov(F_i, F_j) &= E(F_i F_j^t) \\
&= E\left(\boldsymbol{\alpha}_i^t \boldsymbol{X}, \left(\boldsymbol{\alpha}_j^t \boldsymbol{X}\right)^t\right) \\
&= \boldsymbol{\alpha}_i^t E(\boldsymbol{X}\boldsymbol{X}^t) \boldsymbol{\alpha}_j \\
&= \boldsymbol{\alpha}_i^t \boldsymbol{\Sigma} \boldsymbol{\alpha}_j \\
&= \boldsymbol{\alpha}_i^t \lambda_j \boldsymbol{\alpha}_j \\
&= \lambda_j \boldsymbol{\alpha}_i^t \boldsymbol{\alpha}_j = 0 \qquad (2.1.2)
\end{aligned}
$$

Berdasarkan (2.1.2), $cov(F_i, F_j) = 0$, maka dapat disimpulkan bahwa F_1, F_2, …, F_p, tidak berkorelasi satu dengan yang lain.

2.1.2 Penentuan Jumlah Komponen Utama yang Akan Diperhitungkan dalam Analisis.

Telah disebutkan sebelumnya bahwa tujuan dari penggunaan metode analisis komponen utama adalah mereduksi banyaknya variabel asal (yaitu, X_1, X_2, …, X_p) menjadi sejumlah komponen utama yang tidak saling berkorelasi, dan mencakup sebanyak mungkin variansi dari X_1, X_2, …, X_p. Karena banyaknya komponen utama maksimum yang bisa dibentuk adalah sebanyak p, maka banyaknya komponen utama yang akan diperhitungkan dalam analisis adalah kurang dari p. Salah satu kriteria yang bisa digunakan untuk menentukan banyaknya komponen utama yang akan diperhitungkan dalam analisis adalah persentasi dari total variansi.

Sebelumnya, akan dibuktikan bahwa total variansi dari X_1, X_2, …, X_p sama dengan total variansi dari F_1, F_2, …, F_p.

Bukti :

$\boldsymbol{\Sigma}$ merupakan matriks variansi-kovariansi yang dari X_1, X_2, …, X_p. Sebut, $\boldsymbol{A}^t = (\boldsymbol{\alpha_1} \quad \boldsymbol{\alpha_2} \quad \cdots \quad \boldsymbol{\alpha}_p)$. Karena $\boldsymbol{\alpha}_i^t\boldsymbol{\alpha}_j = 0$, dan $\boldsymbol{\alpha}_i^t\boldsymbol{\alpha}_i = 1$, maka $\boldsymbol{A^tA} = \boldsymbol{I} = \boldsymbol{AA^t}$. $\boldsymbol{\Sigma}$ merupakan matriks yang simetri, sehingga $\boldsymbol{\Sigma}$ dapat didiagonalisasi menjadi

$$\boldsymbol{A^t\Sigma A} = \boldsymbol{D} \text{ dan } \boldsymbol{\Sigma} = \boldsymbol{ADA^t} \qquad (2.1.3)$$

D adalah matriks diagonal dengan elemen diagonalnya adalah nilai eigen dari matriks $\boldsymbol{\Sigma}$, yaitu , $\lambda_1, \lambda_2, \ldots, \lambda_p$.

Sebut, $\boldsymbol{F^t} = (F_\mathbf{1} \quad F_\mathbf{2} \quad \cdots \quad F_p)$.

$$\boldsymbol{F} = \begin{pmatrix} F_1 \\ F_2 \\ \vdots \\ F_p \end{pmatrix} = \begin{pmatrix} \boldsymbol{\alpha_1^t X} \\ \boldsymbol{\alpha_2^t X} \\ \vdots \\ \boldsymbol{\alpha_3^t X} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\alpha_1^t} \\ \boldsymbol{\alpha_2^t} \\ \vdots \\ \boldsymbol{\alpha_3^t} \end{pmatrix} \boldsymbol{X} = \boldsymbol{A^tX} \qquad (2.1.4)$$

Misalkan, $\boldsymbol{\Sigma_F}$ adalah matriks variansi-kovariansi dari F_1, F_2, …, F_p, maka $\boldsymbol{\Sigma_F}$ dapat ditulis sebagai berikut

$$\boldsymbol{\Sigma_F} = E(\boldsymbol{FF^t})$$

$$= E(\boldsymbol{A}^t\boldsymbol{X}(\boldsymbol{A}^t\boldsymbol{X})^t)$$

$$= \boldsymbol{A}^t E(\boldsymbol{X}\boldsymbol{X}^t)\boldsymbol{A}$$

$$= \boldsymbol{A}^t\boldsymbol{\Sigma}\boldsymbol{A}$$

$$= \boldsymbol{D} \tag{2.1.5}$$

Total variansi dari $X_1, X_2, \ldots, X_p$ dapat diperoleh dari $\boldsymbol{tr}(\boldsymbol{\Sigma}) = \sum_{i=1}^{p} \sigma_{ii}$, di mana σ_{ii} adalah variansi dari X_i untuk setiap i = 1, 2, …, p. Total variansi dari $F_1, F_2, \ldots, F_p$ dapat diperoleh dari $\boldsymbol{tr}(\boldsymbol{\Sigma_F}) = \boldsymbol{tr}(\boldsymbol{D}) = \sum_{i=1}^{p} \lambda_i$.

$$\boldsymbol{tr}(\boldsymbol{\Sigma}) = \boldsymbol{tr}(\boldsymbol{A}\boldsymbol{D}\boldsymbol{A}^t)$$

$$= \boldsymbol{tr}(\boldsymbol{A}^t\boldsymbol{A}\boldsymbol{D})$$

$$= \boldsymbol{tr}(\boldsymbol{D})$$

$$= \sum_{i=1}^{p} \lambda_i \tag{2.1.6}$$

Berdasarkan (2.1.6), $(\boldsymbol{\Sigma}) = \sum_{i=1}^{p} \lambda_i$, maka dapat disimpulkan bahwa total variansi dari $X_1, X_2, \ldots, X_p$ sama dengan total variansi dari $F_1, F_2, \ldots, F_p$.

Jika p komponen utama diperhitungkan dalam analisis, maka seluruh variansi dari $X_1, X_2, \ldots, X_p$ tercakup dalam $F_1, F_2, \ldots, F_p$. Tetapi, karena $var(F_1) \geq var(F_2) \geq \cdots \geq var(F_p)$, maka dapat dipilih m komponen utama yang pertama, di mana $m < p$, sedemikian sehingga m komponen utama tersebut mencakup sebanyak mungkin total variansi dari $X_1, X_2, \ldots, X_p$. Kriteria yang digunakan adalah persentasi dari total variansi yang ditentukan oleh peneliti. Persentasi total variansi dari m komponen utama yang pertama dapat diperoleh sebagai berikut.

$$\textit{Persentasi total variansi dari m komponen utama yang pertama}$$
$$= \frac{var(F_1) + var(F_2) + \cdots + var(F_m)}{var(F_1) + var(F_2) + \cdots + var(F_p)} \times 100\%$$

$$= \frac{\lambda_1 + \lambda_2 + \cdots + \lambda_m}{\lambda_1 + \lambda_2 + \cdots + \lambda_p} \times 100\%$$

2.2 *Optimal Scaling*

Dalam penelitian, khususnya penelitian sosial, seringkali data kategorik lebih mudah didapatkan dibandingkan data numerik. Namun, sebagian besar metode statistik yang digunakan dalam analisis data memerlukan data numerik. Salah satu metode statistik yang memerlukan data numerik untuk setiap variabelnya adalah Analisis Komponen Utama. Jika variabel yang dilibatkan dalam analisis adalah variabel kategorik, maka akan diusahakan untuk memberikan nilai numerik kepada setiap kategori untuk setiap variabel, yang disebut kuantifikasi. Karena setiap pengamatan termasuk pada kategori tertentu dari variabel kategorik, dan dengan kuantifikasi, setiap kategori diberikan nilai numerik, maka setiap pengamatan akan memiliki nilai numerik. Dengan perkataan lain, variabel kategorik yang telah dikuantifikasi dapat dianggap sebagai variabel numerik. Dengan demikian, metode statistik yang memerlukan data numerik tetap bisa digunakan untuk analisis data.

Kuantifikasi data kategorik sangat mempengaruhi hasil analisis, karena kuantifikasi yang berbeda akan memberikan hasil analisis yang berbeda. Oleh karena itu, diperlukan suatu teknik statistik yang bisa digunakan untuk kuantifikasi yang tepat. Salah satu teknik statistik yang banyak digunakan adalah *optimal scaling*. *Optimal scaling* adalah suatu proses iterasi untuk mencari solusi pengoptimalan suatu fungsi.

Telah dijelaskan pada subbab sebelumnya bahwa metode analisis komponen utama mengasumsikan bahwa variabel-variabel yang akan direduksi (disebut variabel asal) menjadi beberapa komponen utama saling *linearity related*, dan pembentukan komponen utama didasarkan pada matriks kovariansi atau korelasi dari variabel-variabel asal, sehingga metode analisis komponen utama tidak bisa digunakan jika variabel-variabel asal merupakan variabel kategorik. Untuk itu, jika variabel-variabel asal merupakan variabel kategorik, maka teknik *optimal scaling* perlu diterapkan di dalam analisis komponen utama. Dalam analisis komponen utama, prinsip optimal scaling diterapkan untuk kuantifikasi yang tepat kepada setiap variabel yang terlibat dalam analisis dengan memaksimumkan total variansi yang dicakup oleh komponen utama. Analisis komponen utama yang menerapkan teknik *optimal scaling* disebut PRINCALS.

BAB 3

PEMBENTUKAN INDIKATOR SASARAN DENGAN PROXY MEANS TEST BERDASARKAN METODE PRINCALS

3.1 PRINCALS

Seperti telah disebutkan dalam bab terdahulu, PRINCALS merupakan pengembangan dari Analisis Komponen Utama, di mana di dalamnya dilakukan kuantifikasi berdasarkan teknik *optimal scaling*.

Misalkan, $X_1, X_2, \ldots, X_p$ merupakan variabel-variabel kategorik dalam analisis komponen utama yang ingin direduksi menjadi beberapa komponen utama. Misalkan, $C_{(j)}$ adalah banyaknya kategori dari variabel X_j, j = 1, 2, …, p. Sebut, x_{cj} adalah kategori ke-c dari variabel X_j yang merupakan ukuran *nonmetric*, di mana c = 1, 2, …, $C_{(j)}$, dan j = 1, 2, …, p.

Pada prinsipnya, di dalam PRINCALS terdapat dua hal yang akan dilakukan, yaitu

1. Melakukan kuantifikasi terhadap x_{cj} menjadi suatu nilai numerik, sebut sebagai q_{cj}, untuk c = 1, 2, …, $C_{(j)}$, dan j = 1, 2, …, p.

   Sebut, $\boldsymbol{v_j} = \begin{pmatrix} q_{1j} \\ q_{2j} \\ \vdots \\ q_{C_{(j)}j} \end{pmatrix}$

   Misalkan, terdapat n pengamatan, di mana setiap pengamatan termasuk pada salah satu kategori tertentu dari variabel X_j, dengan j = 1, 2, …, p. Definisikan suatu fungsi

   g_{ic} = 1, jika pengamatan ke-i berada pada kategori ke-c untuk variabel X_j

   = 0, jika pengamatan ke-i tidak berada pada kategori ke-c untuk variabel X_j

   di mana i = 1, 2, …, n, j = 1, 2, …, p, dan c = 1, 2, …, $C_{(j)}$.

Berdasarkan definisi tersebut, didapatkan suatu matriks indikator G_j berukuran (n x $C_{(j)}$), untuk setiap variabel X_j, di mana elemen matriks G_j adalah { g_{ic}}.

Pandang, $\boldsymbol{q_j} = \boldsymbol{G_j v_j}$, untuk j = 1, 2, …, p.

$\boldsymbol{q_j}$ adalah matriks berukuran (n x 1). Elemen-elemen $\boldsymbol{q_j}$ merupakan nilai numerik dari n pengamatan untuk variabel X_j. Jika elemen-elemen $\boldsymbol{q_j}$ dianggap sebagai nilai dari suatu variabel numerik Q_j untuk n pengamatan, maka Q_j dapat dianggap sebagai variabel numerik yang berkaitan dengan variabel kategorik X_j, di mana j = 1, 2, …, p.

Untuk mempermudah penjelasan dari notasi di atas, diberikan contoh sebagai berikut:

Misalkan, terdapat empat variabel kategorik X_1, X_2, X_3, X_4, di mana X_1 memiliki tiga kategori, yaitu x_{11}, x_{21}, x_{31}, X_2 memiliki dua kategori, yaitu x_{12}, x_{22}, X_3 memiliki tiga kategori, yaitu x_{13}, x_{23}, x_{33}, dan X_4 memiliki empat kategori, yaitu $x_{14}, x_{24}, x_{34}, x_{44}$.

Misalkan, diambil lima pengamatan, dan diperoleh data sebagai berikut:

Tabel 3.1 Contoh Data dengan Lima Pengamatan untuk Empat Variabel Kategorik

	X_1	X_2	X_3	X_4
1.	x_{11}	x_{22}	x_{33}	x_{24}
2.	x_{11}	x_{22}	x_{13}	x_{44}
3.	x_{21}	x_{12}	x_{33}	x_{14}
4.	x_{31}	x_{12}	x_{23}	x_{34}
5.	x_{11}	x_{22}	x_{13}	x_{34}

Berdasarkan lima pengamatan di atas, diperoleh matriks indikator untuk masing-masing X_j, di mana j = 1, 2, 3, 4, sebagai berikut:

$$G_1 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \quad G_2 = \begin{pmatrix} 0 & 1 \\ 0 & 1 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \quad G_3 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

$$G_4 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Misalkan, untuk masing-masing kategori dari variabel X_1, diberikan nilai numerik q_{11}, q_{21}, q_{31}. Untuk masing-masing kategori dari variabel X_2, diberikan nilai numerik q_{12}, q_{22}. Untuk masing-masing kategori dari variabel X_3, diberikan nilai numerik q_{13}, q_{23}, q_{33}. Untuk masing-masing kategori dari variabel X_4, diberikan nilai numerik $q_{14}, q_{24}, q_{34}, q_{44}$. Dengan kuantifikasi tersebut, diperoleh vektor $\boldsymbol{v_j}$, j = 1, 2, 3, 4, sebagai berikut:

$$\boldsymbol{v_1} = \begin{pmatrix} q_{11} \\ q_{21} \\ q_{31} \end{pmatrix} \quad \boldsymbol{v_2} = \begin{pmatrix} q_{12} \\ q_{22} \end{pmatrix} \quad \boldsymbol{v_3} = \begin{pmatrix} q_{13} \\ q_{23} \\ q_{33} \end{pmatrix} \quad \boldsymbol{v_4} = \begin{pmatrix} q_{14} \\ q_{24} \\ q_{34} \\ q_{44} \end{pmatrix}$$

Vektor $\boldsymbol{q_j}$, untuk j = 1, 2, 3, 4, diperoleh dengan

$$\boldsymbol{q_1} = \boldsymbol{G_1 v_1} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} q_{11} \\ q_{21} \\ q_{31} \end{pmatrix} = \begin{pmatrix} q_{11} \\ q_{11} \\ q_{21} \\ q_{31} \\ q_{11} \end{pmatrix}$$

$$\boldsymbol{q_2} = \boldsymbol{G_2 v_2} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} q_{12} \\ q_{22} \end{pmatrix} = \begin{pmatrix} q_{22} \\ q_{22} \\ q_{12} \\ q_{12} \\ q_{22} \end{pmatrix}$$

$$q_3 = G_3 v_3 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} q_{13} \\ q_{23} \\ q_{33} \end{pmatrix} = \begin{pmatrix} q_{33} \\ q_{13} \\ q_{33} \\ q_{23} \\ q_{13} \end{pmatrix}$$

$$q_4 = G_4 v_4 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} q_{14} \\ q_{24} \\ q_{34} \\ q_{44} \end{pmatrix} = \begin{pmatrix} q_{24} \\ q_{44} \\ q_{14} \\ q_{34} \\ q_{34} \end{pmatrix}$$

Berikut ini adalah n pengamatan untuk variabel X_1, X_2, X_3, X_4 yang telah diubah menjadi data numerik Q_1, Q_2, Q_3, Q_4.

Tabel 3.2 Contoh Data dengan Lima Pengamatan untuk Empat Variabel Kategorik yang Telah Dikuantifikasi

	Q_1	Q_2	Q_3	Q_4
1.	q_{11}	q_{22}	q_{33}	q_{24}
2.	q_{11}	q_{22}	q_{13}	q_{44}
3.	q_{21}	q_{12}	q_{33}	q_{14}
4.	q_{31}	q_{12}	q_{23}	q_{34}
5.	q_{11}	q_{22}	q_{13}	q_{34}

2. Mereduksi Q_j, j = 1, 2, …, p, menjadi m komponen utama, sebut sebagai $F_1, F_2, \dots, F_m$, di mana $m \le p$ yang mencakup semaksimal mungkin total variansi dari $Q_1, Q_2, \dots, Q_p$.

Yang menjadi permasalahan dalam PRINCALS adalah mencari nilai q_{cj} yang optimal (paling tepat), sebut sebagai q_{cj}^*, dan m komponen utama yang didapat berdasarkan nilai q_{cj}^*. Hal ini dilakukan secara iteratif dan simultan berdasarkan butir (1) dan (2) yang telah disebutkan sebelumnya sehingga memaksimumkan total variansi dari $Q_1, Q_2, \dots, Q_p$ yang dicakup oleh komponen-

komponen utama $F_1, F_2, \dots, F_m$. Iterasi tersebut dilakukan dengan dua tahapan, yaitu tahap inisialisasi dan tahap iterasi.

3.1.1 Tahap Inisialisasi

Pada tahap inisialisasi, diberikan nilai numerik awal terhadap kategori x_{cj} untuk setiap variabel X_j, yang disebut $\boldsymbol{v_j}^{(\mathbf{0})}$, di mana j = 1, 2, …, p, dan c = 1, 2, …, $C_{(j)}$. Berdasarkan $\boldsymbol{v_j}^{(\mathbf{0})}$, bentuk $\boldsymbol{q_j}^{(\mathbf{0})} = \boldsymbol{G_j}\boldsymbol{v_j}^{(\mathbf{0})}$, untuk j = 1, 2, …, p. Selanjutnya, berdasarkan elemen-elemen dari $\boldsymbol{q_1}^{(\mathbf{0})}, \boldsymbol{q_2}^{(\mathbf{0})}, \dots, \boldsymbol{q_p}^{(\mathbf{0})}$, yang merupakan data numerik dari variabel $Q_1, Q_2, \dots, Q_p$, akan dibentuk m komponen utama awal, yang disebut sebagai $F_1^{(0)}, F_2^{(0)}, \dots, F_m^{(0)}$. Nilai dari $F_1^{(0)}, F_2^{(0)}, \dots, F_m^{(0)}$ dan $\boldsymbol{q_1}^{(\mathbf{0})}, \boldsymbol{q_2}^{(\mathbf{0})}, \dots, \boldsymbol{q_p}^{(\mathbf{0})}$ akan menjadi nilai awal untuk tahap iterasi.

Karena tidak ada aturan dalam pemberian nilai numerik awal, maka $\boldsymbol{q_1}^{(\mathbf{0})}, \boldsymbol{q_2}^{(\mathbf{0})}, \dots, \boldsymbol{q_p}^{(\mathbf{0})}$ perlu distandarisasi. Misalkan, $\boldsymbol{q_j}^{(\mathbf{0})}*$ adalah vektor yang elemen-elemennya merupakan elemen $\boldsymbol{q_j}^{(\mathbf{0})}$ yang telah distandarisasi.

Sebut, $\boldsymbol{Q}^{(\mathbf{0})} * = \begin{pmatrix} \boldsymbol{q_1}^{(\mathbf{0})} * & \boldsymbol{q_2}^{(\mathbf{0})} * & \dots & \boldsymbol{q_p}^{(\mathbf{0})} * \end{pmatrix}$

Misalkan, untuk contoh sebelumnya, kategori-kategori dari variabel X_1, X_2, X_3, X_4 diberikan nilai numerik awal sebagai berikut:

Untuk variabel X_1, misalkan, $x_{11} = q_{11}^{(0)} = 1$, $x_{21} = q_{21}^{(0)} = 2$,

$$x_{31} = q_{31}^{(0)} = 3$$

Untuk variabel X_2, misalkan, $x_{12} = q_{12}^{(0)} = 1$, $x_{22} = q_{22}^{(0)} = 2$

Untuk variabel X_3, misalkan, $x_{13} = q_{13}^{(0)} = 1$, $x_{23} = q_{23}^{(0)} = 2$,

$$x_{33} = q_{33}^{(0)} = 3$$

Untuk variabel X_3, misalkan, $x_{14} = q_{14}^{(0)} = 1$, $x_{24} = q_{24}^{(0)} = 2$,

$$x_{34} = q_{34}^{(0)} = 3,\ x_{44} = q_{44}^{(0)} = 4$$

Dengan pemberian nilai numerik awal terhadap setiap kategori dari variabel X_1, X_2, X_3, X_4, diperoleh $\boldsymbol{v_j}^{(0)}$, j = 1, 2, 3, 4, sebagai berikut:

$$\boldsymbol{v_1}^{(0)} = \begin{pmatrix}1\\2\\3\end{pmatrix} \quad \boldsymbol{v_2}^{(0)} = \begin{pmatrix}1\\2\end{pmatrix} \quad \boldsymbol{v_3}^{(0)} = \begin{pmatrix}1\\2\\3\end{pmatrix} \quad \boldsymbol{v_4}^{(0)} = \begin{pmatrix}1\\2\\3\\4\end{pmatrix}$$

Selanjutnya, dibentuk $\boldsymbol{q_j}^{(0)}$, untuk j = 1, 2, 3, 4 sebagai berikut

$$\boldsymbol{q_1}^{(0)} = \boldsymbol{G_1 v_1}^{(0)} = \begin{pmatrix}1\\1\\2\\3\\1\end{pmatrix} \qquad \boldsymbol{q_2}^{(0)} = \boldsymbol{G_2 v_2}^{(0)} = \begin{pmatrix}2\\2\\1\\1\\2\end{pmatrix}$$

$$\boldsymbol{q_3}^{(0)} = \boldsymbol{G_3 v_3}^{(0)} = \begin{pmatrix}3\\1\\3\\2\\1\end{pmatrix} \qquad \boldsymbol{q_4}^{(0)} = \boldsymbol{G_4 v_4}^{(0)} = \begin{pmatrix}2\\4\\1\\3\\4\end{pmatrix}$$

Berikut ini adalah n pengamatan untuk variabel X_1, X_2, X_3, X_4 yang telah diubah menjadi data numerik, dengan melakukan kuantifikasi awal kepada setiap kategori dari X_1, X_2, X_3, X_4, yang disebut Q_1, Q_2, Q_3, Q_4.

Tabel 3.3 Contoh Data dengan Lima Pengamatan untuk Empat Variabel Kategorik yang Telah Diberikan Nilai Awal

	Q_1	Q_2	Q_3	Q_4
1.	1	2	3	2
2.	1	2	1	4
3.	2	1	3	1
4.	3	1	2	3
5.	1	2	1	4
Mean	1.6	1.6	2	1.7
S.D	0.894427	0.547723	1	1.30384

Vektor $\boldsymbol{q_j}^{(\mathbf{0})}*$ untuk setiap $\boldsymbol{q_j}^{(\mathbf{0})}$, j = 1, 2, 3, 4, adalah sebagai berikut

$$\boldsymbol{q_1}^{(\mathbf{0})}*=\begin{pmatrix}-0.671\\-0.671\\0.447\\1.565\\-0.671\end{pmatrix}\quad \boldsymbol{q_2}^{(\mathbf{0})}*=\begin{pmatrix}0.73\\0.73\\-1.095\\-1.095\\0.73\end{pmatrix}$$

$$\boldsymbol{q_3}^{(\mathbf{0})}*=\begin{pmatrix}1\\-1\\1\\0\\-1\end{pmatrix}\qquad \boldsymbol{q_4}^{(\mathbf{0})}*=\begin{pmatrix}-0.614\\0.92\\-1.38\\0.153\\0.92\end{pmatrix}$$

Sehingga, matriks $\boldsymbol{Q}^{(\mathbf{0})}$ adalah sebagai berikut:

$$\boldsymbol{Q}^{(\mathbf{0})}=\begin{pmatrix}-0.671 & 0.73 & 1 & -0.614\\-0.671 & 0.73 & -1 & 0.92\\0.447 & -1.095 & 1 & -1.38\\1.565 & -1.095 & 0 & 0.153\\-0.671 & 0.73 & -1 & 0.92\end{pmatrix}$$

Misalkan, $\mathbf{R}^{(\mathbf{0})}$ adalah matriks korelasi dari $Q_1, Q_2, \dots, Q_p$, yang dicari berdasarkan data numerik yang merupakan elemen-elemen $\boldsymbol{q_1}^{(\mathbf{0})}*, \boldsymbol{q_2}^{(\mathbf{0})}*, \dots, \boldsymbol{q_p}^{(\mathbf{0})}*$, dan dapat dibuktikan bahwa $\boldsymbol{R}^{(\mathbf{0})}=\frac{\mathbf{1}}{\boldsymbol{n}-\mathbf{1}}\boldsymbol{Q}^{(\mathbf{0})^t}\boldsymbol{Q}^{(\mathbf{0})}$ (lampiran).

Misalkan, $l_1^{(0)}, l_2^{(0)}, \dots, l_p^{(0)}$ adalah nilai eigen dari $\boldsymbol{R}^{(\mathbf{0})}$, di mana $l_1^{(0)} \geq l_2^{(0)} \geq \cdots \geq l_p^{(0)}$. Misalkan, $\boldsymbol{\gamma}_1^{(0)}, \boldsymbol{\gamma}_2^{(0)}, \dots, \boldsymbol{\gamma}_p^{(0)}$ merupakan vektor eigen dari matriks $\boldsymbol{R}^{(\mathbf{0})}$, yang bersesuaian dengan nilai eigen $l_1^{(0)}, l_2^{(0)}, \dots, l_p^{(0)}$. Seperti telah disebutkan pada bab sebelumnya, akan dipilih m vektor eigen yang pertama dari matriks korelasi $\boldsymbol{R}^{(\mathbf{0})}$, yaitu $\boldsymbol{\gamma}_1^{(0)}, \boldsymbol{\gamma}_2^{(0)}, \dots, \boldsymbol{\gamma}_m^{(0)}$, yang bersesuaian dengan m nilai eigen yang pertama, yaitu $l_1^{(0)}, l_2^{(0)}, \dots, l_m^{(0)}$, untuk membentuk m komponen utama awal, sebut sebagai $F_1^{(0)}, F_2^{(0)}, \dots, F_m^{(0)}$. Misalkan, $\boldsymbol{f}_k^{(0)}$ adalah nilai dari komponen utama $F_k^{(0)}$ untuk n pengamatan, di mana k = 1, 2, …, m. Berdasarkan metode analisis komponen utama pada bab sebelumnya, $\boldsymbol{f}_k^{(0)}$ diperoleh dengan $\boldsymbol{f}_k^{(0)}=\boldsymbol{Q}^{(\mathbf{0})}\boldsymbol{\gamma}_k^{(0)}$, untuk k =1, 2, …, m.

Misalkan, untuk contoh sebelumnya, Q_1, Q_2, Q_3, Q_4, akan direduksi menjadi dua komponen utama, berdasarkan data numerik yang merupakan elemen-elemen dari $\boldsymbol{q_1}^{(0)}*, \boldsymbol{q_2}^{(0)}*, \boldsymbol{q_3}^{(0)}*, \boldsymbol{q_4}^{(0)}*$. Berdasarkan matriks korelasi $\boldsymbol{R}^{(0)} = \frac{1}{4}\boldsymbol{Q}^{(0)^t}\boldsymbol{Q}^{(0)}$, diperoleh dua nilai eigen yang pertama, yaitu $l_1^{(0)} = 2.747$, dan $l_2^{(0)} = 1.1629$, yang bersesuaian dengan vektor eigen

$$\boldsymbol{\gamma}_1^{(0)} = \begin{pmatrix} -0.448 \\ 0.532 \\ -0.496 \\ 0.52 \end{pmatrix} \qquad \boldsymbol{\gamma}_2^{(0)} = \begin{pmatrix} 0.604 \\ -0.411 \\ -0.507 \\ 0.458 \end{pmatrix}$$

Nilai dari komponen utama $F_1^{(0)}$ dan $F_2^{(0)}$, yaitu $\boldsymbol{f}_1^{(0)}$, dan $\boldsymbol{f}_2^{0}$, untuk lima pengamatan adalah

$$\boldsymbol{f}_1^{(0)} = \boldsymbol{Q}^{(0)}\boldsymbol{\gamma}_1^{(0)} = \begin{pmatrix} -0.671 & 0.73 & 1 & -0.614 \\ -0.671 & 0.73 & -1 & 0.92 \\ 0.447 & -1.095 & 1 & -1.38 \\ 1.565 & -1.095 & 0 & 0.153 \\ -0.671 & 0.73 & -1 & 0.92 \end{pmatrix} \begin{pmatrix} -0.448 \\ 0.532 \\ -0.496 \\ 0.52 \end{pmatrix} = \begin{pmatrix} -0.126 \\ 1.664 \\ -1.996 \\ -1.205 \\ 1.664 \end{pmatrix}$$

$$\boldsymbol{f}_2^{(0)} = \boldsymbol{Q}^{(0)}\boldsymbol{\gamma}_2^{(0)} = \begin{pmatrix} -0.671 & 0.73 & 1 & -0.614 \\ -0.671 & 0.73 & -1 & 0.92 \\ 0.447 & -1.095 & 1 & -1.38 \\ 1.565 & -1.095 & 0 & 0.153 \\ -0.671 & 0.73 & -1 & 0.92 \end{pmatrix} \begin{pmatrix} 0.604 \\ -0.411 \\ -0.507 \\ 0.458 \end{pmatrix} = \begin{pmatrix} -1.493 \\ 0.223 \\ -0.418 \\ 1.466 \\ 0.223 \end{pmatrix}$$

$\boldsymbol{f}_1^{(0)}, \boldsymbol{f}_2^{(0)}, \dots, \boldsymbol{f}_m^{(0)}$, yang merupakan nilai dari $F_1^{(0)}, F_2^{(0)}, \dots, F_m^{(0)}$, dan $\boldsymbol{q}_1^{(0)}, \boldsymbol{q}_2^{(0)}, \dots, \boldsymbol{q}_m^{(0)}$ menjadi nilai awal untuk tahap iterasi.

3.1.2 Tahap Iterasi

Seperti telah dijelaskan pada subbab sebelumnya bahwa $\boldsymbol{f}_1^{(0)}, \boldsymbol{f}_2^{(0)}, \dots, \boldsymbol{f}_m^{(0)}$, yang merupakan nilai dari $F_1^{(0)}, F_2^{(0)}, \dots, F_m^{(0)}$, dan $\boldsymbol{q}_1^{(0)}, \boldsymbol{q}_2^{(0)}, \dots, \boldsymbol{q}_p^{(0)}$ menjadi nilai awal untuk tahap iterasi. Telah dibuktikan pada bab sebelumnya bahwa $F_1^{(0)}, F_2^{(0)}, \dots, F_m^{(0)}$ tidak saling berkorelasi atau $\boldsymbol{f}_1^{(0)}, \boldsymbol{f}_2^{(0)}, \dots, \boldsymbol{f}_m^{(0)}$ saling ortogonal.

Pandang, setiap variabel numerik Q_j yang berkaitan dengan variabel kategorik X_j, di mana j = 1, 2, …, p, sebagai kombinasi linier dari m komponen utama $F_1, F_2, \dots, F_m$.

$$Q_j = a_{j1}F_1 + a_{j2}F_2 + \cdots + a_{jm}\, F_m \qquad (3.2.1)$$

di mana $F_1, F_2, \dots, F_m$ tidak saling berkorelasi satu dengan yang lain, atau dengan perkataan lain $cov(F_k, F_{k'}) = 0$, untuk $k \neq k'$, k = 1, 2, …, m.

Misalkan, $\boldsymbol{q_j}$ adalah nilai numerik dari variabel Q_j untuk n pengamatan, di mana j = 1, 2, …, p, dan $\boldsymbol{f_k}$ adalah nilai dari komponen utama F_k untuk n pengamatan, di mana k = 1, 2, …, m, maka berdasarkan (3.2.1), untuk setiap j = 1, 2, …, p, maka $\boldsymbol{q_j}$ dapat dipandang sebagai kombinasi linier dari $\boldsymbol{f_1}, \boldsymbol{f_2}, \dots, \boldsymbol{f_m}$.

$$\boldsymbol{q_j} = a_{j1}\boldsymbol{f_1} + a_{j2}\boldsymbol{f_2} + \cdots + a_{jm}\, \boldsymbol{f_m} \qquad (3.2.2)$$

di mana kovariansi antara F_k dan $F_{k'}$ berdasarkan elemen-elemen dari $\boldsymbol{f_k}$ dan $\boldsymbol{f_{k'}}$ harus memenuhi $\frac{1}{n}\boldsymbol{f_k}^t\boldsymbol{f_{k'}} = \frac{1}{n}\sum_{i=1}^{n} f_{ik}f_{ik'} = 0$, atau $\boldsymbol{f_k}^t\boldsymbol{f_{k'}} = 0$ untuk $k \neq k'$.

Sebut, $\boldsymbol{a_j} = \begin{pmatrix} a_{j1} \\ a_{j2} \\ \vdots \\ a_{jm} \end{pmatrix}$ untuk j = 1, 2, …, p.

Sebut, $\boldsymbol{F} = (\boldsymbol{f_1} \quad \boldsymbol{f_2} \quad \cdots \quad \boldsymbol{f_m})$, maka persamaan (3.2.2) dapat ditulis sebagai

$$\boldsymbol{q_j} = a_{j1}\boldsymbol{f_1} + a_{j2}\boldsymbol{f_2} + \cdots + a_{jm}\, \boldsymbol{f_m} = \boldsymbol{F}\boldsymbol{a_j} \qquad (3.2.3)$$

Karena $(m < p)$, maka nilai dari $\boldsymbol{a_j}$ tidak unik.

Tahap iterasi diawali dengan mencari $\boldsymbol{a_j}$, untuk setiap j = 1, 2, …, p, menggunakan $\boldsymbol{f_1}, \boldsymbol{f_2}, \dots, \boldsymbol{f_m}$, dan $\boldsymbol{q_1}, \boldsymbol{q_2}, \dots, \boldsymbol{q_p}$, dengan meminimumkan fungsi

$$L_j = \left\|\boldsymbol{q_j} - \boldsymbol{F}\boldsymbol{a_j}\right\|^2 = \left(\boldsymbol{q_j} - \boldsymbol{F}\boldsymbol{a_j}\right)^t\left(\boldsymbol{q_j} - \boldsymbol{F}\boldsymbol{a_j}\right) \qquad (3.2.4)$$

Gifi (1990) telah membuktikan bahwa meminimumkan fungsi L_j pada (3.2.4) sama dengan meminimumkan fungsi

$$Z_j = tr\left[\left(\boldsymbol{q}_j\boldsymbol{\alpha}_j{}^t - \boldsymbol{F}\right)^t\left(\boldsymbol{q}_j\boldsymbol{\alpha}_j{}^t - \boldsymbol{F}\right)\right] \tag{3.2.5}$$

untuk j = 1, 2, …, p. Z_j pada persamaan (3.2.5) dapat ditulis kembali sebagai berikut

$$\begin{aligned} Z_j &= tr\left[\left(\boldsymbol{q}_j\boldsymbol{\alpha}_j{}^t - \boldsymbol{F}\right)^t\left(\boldsymbol{q}_j\boldsymbol{\alpha}_j{}^t - \boldsymbol{F}\right)\right] \\ &= tr\left[\boldsymbol{\alpha}_j\boldsymbol{q}_j{}^t\boldsymbol{q}_j\boldsymbol{\alpha}_j{}^t - \boldsymbol{\alpha}_j\boldsymbol{q}_j{}^t\boldsymbol{F} - \boldsymbol{F}^t\boldsymbol{q}_j\boldsymbol{\alpha}_j{}^t + \boldsymbol{F}^t\boldsymbol{F}\right] \\ &= tr\left[\boldsymbol{\alpha}_j\boldsymbol{q}_j{}^t\boldsymbol{q}_j\boldsymbol{\alpha}_j{}^t\right] - tr\left[\boldsymbol{\alpha}_j\boldsymbol{q}_j{}^t\boldsymbol{F}\right] - tr\left[\boldsymbol{F}^t\boldsymbol{q}_j\boldsymbol{\alpha}_j{}^t\right] + tr\left[\boldsymbol{F}^t\boldsymbol{F}\right] \end{aligned} \tag{3.2.6}$$

Karena $tr\left[\boldsymbol{\alpha}_j\boldsymbol{q}_j{}^t\boldsymbol{F}\right] = tr\left[\left(\boldsymbol{\alpha}_j\boldsymbol{q}_j{}^t\boldsymbol{F}\right)^t\right] = tr\left[\boldsymbol{F}^t\boldsymbol{q}_j\boldsymbol{\alpha}_j{}^t\right]$ maka

$$= tr\left[\boldsymbol{\alpha}_j\boldsymbol{q}_j{}^t\boldsymbol{q}_j\boldsymbol{\alpha}_j{}^t\right] - tr\left[\boldsymbol{\alpha}_j\boldsymbol{q}_j{}^t\boldsymbol{F}\right] - tr\left[\boldsymbol{F}^t\boldsymbol{q}_j\boldsymbol{\alpha}_j{}^t\right] + tr\left[\boldsymbol{F}^t\boldsymbol{F}\right]$$

$$= tr\left[\boldsymbol{\alpha}_j\boldsymbol{q}_j{}^t\boldsymbol{q}_j\boldsymbol{\alpha}_j{}^t\right] - 2tr\left[\boldsymbol{\alpha}_j\boldsymbol{q}_j{}^t\boldsymbol{F}\right] + tr\left[\boldsymbol{F}^t\boldsymbol{F}\right] \tag{3.2.7}$$

Nilai $\boldsymbol{\alpha}_j$ dicari dengan meminimumkan Z_j pada persamaan (3.2.7) yang sama seperti meminimumkan L_j pada persamaan (3.2.4) untuk j = 1, 2, …, p.

$$\frac{\partial}{\partial \alpha_j} Z_j = 0 \tag{3.2.8}$$

$$\frac{\partial}{\partial \alpha_j}\left(tr\left[\boldsymbol{\alpha}_j\boldsymbol{q}_j{}^t\boldsymbol{q}_j\boldsymbol{\alpha}_j{}^t\right] - 2tr\left[\boldsymbol{\alpha}_j\boldsymbol{q}_j{}^t\boldsymbol{F}\right] + tr\left[\boldsymbol{F}^t\boldsymbol{F}\right]\right) = 0 \tag{3.2.9}$$

$$2\left(\boldsymbol{q}_j{}^t\boldsymbol{q}_j\boldsymbol{\alpha}_j{}^t\right)^t - 2\left(\boldsymbol{q}_j{}^t\boldsymbol{F}\right)^t = \boldsymbol{0} \tag{3.2.10}$$

Maka, $\boldsymbol{\alpha}_j$, yang meminimumkan Z_j, untuk setiap j = 1, 2, …, p, harus memenuhi

$$\boldsymbol{\alpha}_j\boldsymbol{q}_j{}^t\boldsymbol{q}_j = \boldsymbol{F}^t\boldsymbol{q}_j$$

$$\boldsymbol{\alpha}_j = \boldsymbol{F}^t\boldsymbol{q}_j\left(\boldsymbol{q}_j{}^t\boldsymbol{q}_j\right)^{-\boldsymbol{1}} \tag{3.2.11}$$

Iterasi diawali dengan menggunakan nilai awal $\boldsymbol{f}_1^{(\boldsymbol{0})}, \boldsymbol{f}_2^{(\boldsymbol{0})}, \dots, \boldsymbol{f}_m^{(\boldsymbol{0})}$, dan $\boldsymbol{q}_1^{(\boldsymbol{0})}, \boldsymbol{q}_2^{(\boldsymbol{0})}, \dots, \boldsymbol{q}_m^{(\boldsymbol{0})}$ untuk mencari $\boldsymbol{\alpha}_j$. Misalkan, solusi nilai $\boldsymbol{\alpha}_j$ yang didapat menggunakan nilai awal tersebut disebut $\boldsymbol{\alpha}_j{}^{(\boldsymbol{1})}$. Jadi, nilai $\boldsymbol{\alpha}_j{}^{(\boldsymbol{1})}$ diperoleh dengan

$$\boldsymbol{a}_j^{(1)} = \boldsymbol{F}^{(0)^t}\boldsymbol{q}_j^{(0)}\left(\boldsymbol{q}_j^{(0)^t}\boldsymbol{q}_j^{(0)}\right)^{-1} \qquad (3.2.12)$$

Setelah mendapatkan, $\boldsymbol{a}_j^{(1)}$, tahap iterasi dilanjutkan dengan mencari $\boldsymbol{q}_j^{(1)}$ menggunakan $\boldsymbol{a}_j^{(1)}$, di mana j = 1, 2, …, p, dan $\boldsymbol{f}_1^{(0)}, \boldsymbol{f}_2^{(0)}, \dots, \boldsymbol{f}_m^{(0)}$ yang meminimumkan Z_j pada persamaan (3.2.5).

$$Z_j = tr\left[\left(\boldsymbol{q}_j^{(1)}\boldsymbol{a}_j^{(1)^t} - \boldsymbol{F}^{(0)}\right)^t\left(\boldsymbol{q}_j^{(1)}\boldsymbol{a}_j^{(1)^t} - \boldsymbol{F}^{(0)}\right)\right]$$

$$= tr\left[\boldsymbol{a}_j^{(1)}\boldsymbol{q}_j^{(1)^t}\boldsymbol{q}_j^{(1)}\boldsymbol{a}_j^{(1)^t} - \boldsymbol{a}_j^{(1)}\boldsymbol{q}_j^{(1)^t}\boldsymbol{F}^{(0)} - \boldsymbol{F}^{(0)^t}\boldsymbol{q}_j^{(1)}\boldsymbol{a}_j^{(1)^t} + \boldsymbol{F}^{(0)^t}\boldsymbol{F}^{(0)}\right]$$

$$= tr\left[\boldsymbol{a}_j^{(1)}\boldsymbol{q}_j^{(1)^t}\boldsymbol{q}_j^{(1)}\boldsymbol{a}_j^{(1)^t}\right] - tr\left[\boldsymbol{a}_j^{(1)}\boldsymbol{q}_j^{(1)^t}\boldsymbol{F}^{(0)}\right] - tr\left[\boldsymbol{F}^{(0)^t}\boldsymbol{q}_j^{(1)}\boldsymbol{a}_j^{(1)^t}\right] + tr\left[\boldsymbol{F}^{(0)^t}\boldsymbol{F}^{(0)}\right] \qquad (3.2.13)$$

Karena $tr\left[\boldsymbol{a}_j^{(1)}\boldsymbol{q}_j^{(1)^t}\boldsymbol{F}^{(0)}\right] = tr\left[\left(\boldsymbol{a}_j^{(1)}\boldsymbol{q}_j^{(1)^t}\boldsymbol{F}^{(0)}\right)^t\right] = tr\left[\boldsymbol{F}^{(0)^t}\boldsymbol{q}_j^{(1)}\boldsymbol{a}_j^{(1)^t}\right]$ maka

$$= tr\left[\boldsymbol{a}_j^{(1)}\boldsymbol{q}_j^{(1)^t}\boldsymbol{q}_j^{(1)}\boldsymbol{a}_j^{(1)^t}\right] - tr\left[\boldsymbol{a}_j^{(1)}\boldsymbol{q}_j^{(1)^t}\boldsymbol{F}^{(0)}\right] - tr\left[\boldsymbol{F}^{(0)^t}\boldsymbol{q}_j^{(1)}\boldsymbol{a}_j^{(1)^t}\right] + tr\left[\boldsymbol{F}^{(0)^t}\boldsymbol{F}^{(0)}\right]$$

$$= tr\left[\boldsymbol{a}_j^{(1)}\boldsymbol{q}_j^{(1)^t}\boldsymbol{q}_j^{(1)}\boldsymbol{a}_j^{(1)^t}\right] - 2tr\left[\boldsymbol{a}_j^{(1)}\boldsymbol{q}_j^{(1)^t}\boldsymbol{F}^{(0)}\right] + tr\left[\boldsymbol{F}^{(0)^t}\boldsymbol{F}^{(0)}\right] \qquad (3.2.14)$$

Karena $\boldsymbol{q}_j^{(1)} = \boldsymbol{G}_j\boldsymbol{v}_j^{(1)}$, maka Z_j pada (3.2.14) dapat ditulis sebagai berikut

$$= tr\left[\boldsymbol{a}_j^{(1)}\boldsymbol{q}_j^{(1)^t}\boldsymbol{q}_j^{(1)}\boldsymbol{a}_j^{(1)^t}\right] - 2tr\left[\boldsymbol{a}_j^{(1)}\boldsymbol{q}_j^{(1)^t}\boldsymbol{F}^{(0)}\right] + tr\left[\boldsymbol{F}^{(0)^t}\boldsymbol{F}^{(0)}\right]$$

$$= tr\left[\boldsymbol{a}_j^{(1)}\left(\boldsymbol{G}_j\boldsymbol{v}_j^{(1)}\right)^t\left(\boldsymbol{G}_j\boldsymbol{v}_j^{(1)}\right)\boldsymbol{a}_j^{(1)^t}\right] - 2tr\left[\boldsymbol{a}_j^{(1)}\left(\boldsymbol{G}_j\boldsymbol{v}_j^{(1)}\right)^t\boldsymbol{F}^{(0)}\right] + tr\left[\boldsymbol{F}^{(0)^t}\boldsymbol{F}^{(0)}\right]$$

$$= tr\left[\boldsymbol{a}_j^{(1)}\boldsymbol{v}_j^{(1)^t}\boldsymbol{G}_j^t\boldsymbol{G}_j\boldsymbol{v}_j^{(1)}\boldsymbol{a}_j^{(1)^t}\right] - 2tr\left[\boldsymbol{a}_j^{(1)}\boldsymbol{v}_j^{(1)^t}\boldsymbol{G}_j^t\boldsymbol{F}^{(0)}\right] + tr\left[\boldsymbol{F}^{(0)^t}\boldsymbol{F}^{(0)}\right]$$

(3.2.15)

Untuk mendapatkan $\boldsymbol{q}_j^{(1)}$, terlebih dahulu harus dicari $\boldsymbol{v_j}^{(\mathbf{1})}$ yang meminimumkan Z_j pada persamaan (3.2.15). $\boldsymbol{v_j}^{(\mathbf{1})}$ yang meminimumkan Z_j diperoleh dengan

$$\frac{\partial}{\partial \boldsymbol{v_j}} Z_j = 0 \qquad (3.2.16)$$

$$\frac{\partial}{\partial \boldsymbol{v_j}}\left(tr\left[\boldsymbol{a}_j^{(\mathbf{1})} \boldsymbol{v}_j^{(\mathbf{1})^t} \boldsymbol{G}_j^t \boldsymbol{G}_j \boldsymbol{v}_j^{(\mathbf{1})} \boldsymbol{a}_j^{(\mathbf{1})^t}\right] - 2tr\left[\boldsymbol{a}_j^{(\mathbf{1})} \boldsymbol{v}_j^{(\mathbf{1})^t} \boldsymbol{G}_j^t \boldsymbol{F}^{(\mathbf{0})}\right] + \right.$$
$$\left. tr\left[\boldsymbol{F}^{(\mathbf{0})^t} \boldsymbol{F}^{(\mathbf{0})}\right]\right) = 0 \qquad (3.2.17)$$

$$2\boldsymbol{G}_j^t \boldsymbol{G}_j \boldsymbol{v}_j^{(\mathbf{1})} \boldsymbol{a}_j^{(\mathbf{1})^t} \boldsymbol{a}_j^{(\mathbf{1})} - 2\boldsymbol{G}_j^t \boldsymbol{F}^{(\mathbf{0})} \boldsymbol{a}_j^{(\mathbf{1})} = \mathbf{0} \qquad (3.2.18)$$

Maka, $\boldsymbol{v}_j^{(\mathbf{1})}$, yang meminimumkan Z_j, untuk setiap j = 1, 2, …, p, harus memenuhi

$$\boldsymbol{G}_j^t \boldsymbol{G}_j \boldsymbol{v}_j^{(\mathbf{1})} \boldsymbol{a}_j^{(\mathbf{1})^t} \boldsymbol{a}_j^{(\mathbf{1})} = \boldsymbol{G}_j^t \boldsymbol{F}^{(\mathbf{0})} \boldsymbol{a}_j^{(\mathbf{1})}$$

$$\boldsymbol{v}_j^{(\mathbf{1})} \boldsymbol{a}_j^{(\mathbf{1})^t} \boldsymbol{a}_j^{(\mathbf{1})} = \left(\boldsymbol{G}_j^t \boldsymbol{G}_j\right)^{-\mathbf{1}} \boldsymbol{G}_j^t \boldsymbol{F}^{(\mathbf{0})} \boldsymbol{a}_j^{(\mathbf{1})}$$

$$\boldsymbol{v}_j^{(\mathbf{1})} = \left(\boldsymbol{G}_j^t \boldsymbol{G}_j\right)^{-\mathbf{1}} \boldsymbol{G}_j^t \boldsymbol{F}^{(\mathbf{0})} \boldsymbol{a}_j^{(\mathbf{1})} \left(\boldsymbol{a}_j^{(\mathbf{1})^t} \boldsymbol{a}_j^{(\mathbf{1})}\right)^{-\mathbf{1}} \qquad (3.2.19)$$

Setelah mendapatkan $\boldsymbol{v}_j^{(\mathbf{1})}$, maka didapat nilai $\boldsymbol{q}_j^{(\mathbf{1})}$ dengan $\boldsymbol{q}_j^{(\mathbf{1})} = \boldsymbol{G}_j \boldsymbol{v}_j^{(\mathbf{1})}$.

Langkah selanjutnya pada tahap iterasi adalah menentukan $\boldsymbol{f}_1^{(1)}, \boldsymbol{f}_2^{(1)}, \dots, \boldsymbol{f}_m^{(1)}$ yang merupakan nilai dari komponen utama $F_1^{(1)}, F_2^{(1)}, \dots, F_m^{(1)}$ menggunakan $\boldsymbol{a}_1^{(1)}, \boldsymbol{a}_2^{(1)}, \dots, \boldsymbol{a}_p^{(1)}$, dan $\boldsymbol{q}_1^{(1)}, \boldsymbol{q}_2^{(1)}, \dots, \boldsymbol{q}_p^{(1)}$. Nilai dari komponen utama $F_1^{(1)}, F_2^{(1)}, \dots, F_m^{(1)}$ dicari dengan meminimumkan fungsi

$$S = \frac{1}{p} \sum_{j=1}^{p} Z_j$$

$$S = \frac{1}{p} \sum_{j=1}^{p} tr\left[\left(\boldsymbol{q}_j^{(\mathbf{1})} \boldsymbol{a}_j^{(\mathbf{1})^t} - \boldsymbol{F}^{(\mathbf{1})}\right)^t \left(\boldsymbol{q}_j^{(\mathbf{1})} \boldsymbol{a}_j^{(\mathbf{1})^t} - \boldsymbol{F}^{(\mathbf{1})}\right)\right]$$

$$= \frac{1}{p}\sum_{j=1}^{p}\Big(tr\Big[\boldsymbol{a}_j{}^{(\mathbf{1})}\boldsymbol{q}_j^{(1)^{\boldsymbol{t}}}\boldsymbol{q}_j^{(1)}\boldsymbol{a}_j{}^{(\mathbf{1})^{\boldsymbol{t}}} - \boldsymbol{a}_j{}^{(\mathbf{1})}\boldsymbol{q}_j^{(1)^{\boldsymbol{t}}}\boldsymbol{F}^{(\mathbf{1})} - \boldsymbol{F}^{(\mathbf{1})^{\boldsymbol{t}}}\boldsymbol{q}_j^{(1)}\boldsymbol{a}_j{}^{(\mathbf{1})^{\boldsymbol{t}}} + \boldsymbol{F}^{(\mathbf{1})^{\boldsymbol{t}}}\boldsymbol{F}^{(\mathbf{1})}\Big]\Big)$$

$$= \frac{1}{p}\sum_{j=1}^{p}\Big(tr\Big[\boldsymbol{a}_j{}^{(\mathbf{1})}\boldsymbol{q}_j^{(1)^{\boldsymbol{t}}}\boldsymbol{q}_j^{(1)}\boldsymbol{a}_j{}^{(\mathbf{1})^{\boldsymbol{t}}}\Big] - tr\Big[\boldsymbol{a}_j{}^{(\mathbf{1})}\boldsymbol{q}_j^{(1)^{\boldsymbol{t}}}\boldsymbol{F}^{(\mathbf{1})}\Big] - tr\Big[\boldsymbol{F}^{(\mathbf{1})^{\boldsymbol{t}}}\boldsymbol{q}_j^{(1)}\boldsymbol{a}_j{}^{(\mathbf{1})^{\boldsymbol{t}}}\Big] + tr\Big[\boldsymbol{F}^{(\mathbf{1})^{\boldsymbol{t}}}\boldsymbol{F}^{(\mathbf{1})}\Big]\Big) \quad (3.2.20)$$

Karena $tr\Big[\boldsymbol{a}_j{}^{(\mathbf{1})}\boldsymbol{q}_j^{(1)^{\boldsymbol{t}}}\boldsymbol{F}^{(\mathbf{1})}\Big] = tr\Big[\Big(\boldsymbol{a}_j{}^{(\mathbf{1})}\boldsymbol{q}_j^{(1)^{\boldsymbol{t}}}\boldsymbol{F}^{(\mathbf{1})}\Big)^{\boldsymbol{t}}\Big] = tr\Big[\boldsymbol{F}^{(\mathbf{1})^{\boldsymbol{t}}}\boldsymbol{q}_j^{(1)}\boldsymbol{a}_j{}^{(\mathbf{1})^{\boldsymbol{t}}}\Big]$ maka

$$S = \frac{1}{p}\sum_{j=1}^{p}\Big[tr\Big[\boldsymbol{a}_j{}^{(\mathbf{1})}\boldsymbol{q}_j^{(1)^{\boldsymbol{t}}}\boldsymbol{q}_j^{(1)}\boldsymbol{a}_j{}^{(\mathbf{1})^{\boldsymbol{t}}}\Big] - 2tr\Big[\boldsymbol{a}_j{}^{(\mathbf{1})}\boldsymbol{q}_j^{(1)^{\boldsymbol{t}}}\boldsymbol{F}^{(\mathbf{1})}\Big] + tr\Big[\boldsymbol{F}^{(\mathbf{1})^{\boldsymbol{t}}}\boldsymbol{F}^{(\mathbf{1})}\Big]\Big] \quad (3.2.21)$$

Nilai dari komponen utama $F_1, F_2, \dots, F_m$ yang meminimumkan S pada persamaan (3.2.21) diperoleh dengan

$$\frac{\partial}{\partial \boldsymbol{F}} S = 0 \quad (3.2.22)$$

$$\frac{\partial}{\partial \boldsymbol{v_j}}\frac{1}{p}\sum_{j=1}^{p}\Big(tr\Big[\boldsymbol{a}_j{}^{(\mathbf{1})}\boldsymbol{q}_j{}^{(\mathbf{1})^{\boldsymbol{t}}}\boldsymbol{q}_j{}^{(\mathbf{1})}\boldsymbol{a}_j{}^{(\mathbf{1})^{\boldsymbol{t}}}\Big] - 2tr\Big[\boldsymbol{a}_j{}^{(\mathbf{1})}\boldsymbol{q}_j{}^{(\mathbf{1})^{\boldsymbol{t}}}\boldsymbol{F}^{(\mathbf{1})}\Big] + tr\Big[\boldsymbol{F}^{(\mathbf{1})^{\boldsymbol{t}}}\boldsymbol{F}^{(\mathbf{1})}\Big]\Big) = 0 \quad (3.2.23)$$

$$\frac{1}{p}\sum_{j=1}^{p}\Big(-2\Big[\boldsymbol{a}_j{}^{(\mathbf{1})}\boldsymbol{q}_j{}^{(\mathbf{1})^{\boldsymbol{t}}}\Big]^{\boldsymbol{t}} + 2\boldsymbol{F}^{(\mathbf{1})}\Big) = 0 \quad (3.2.24)$$

Maka, nilai dari komponen utama $F_1^{(1)}, F_2^{(1)}, \dots, F_m^{(1)}$ yang meminimumkan S harus memenuhi

$$\frac{1}{p}\sum_{j=1}^{p}\Big(-2\boldsymbol{q}_j{}^{(\mathbf{1})}\boldsymbol{a}_j{}^{(\mathbf{1})^{\boldsymbol{t}}} + 2\boldsymbol{F}^{(\mathbf{1})}\Big) = 0$$

$$\sum_{j=1}^{p}\boldsymbol{F}^{(\mathbf{1})} = \sum_{j=1}^{p}\boldsymbol{q}_j{}^{(\mathbf{1})}\boldsymbol{a}_j{}^{(\mathbf{1})^{\boldsymbol{t}}}$$

$$\boldsymbol{F}^{(\mathbf{1})}* = \frac{1}{p}\sum_{j=1}^{p}\boldsymbol{q}_j{}^{(\mathbf{1})}\boldsymbol{a}_j{}^{(\mathbf{1})^{\boldsymbol{t}}} \quad (3.2.25)$$

Berdasarkan matriks $\boldsymbol{F}^{(\mathbf{1})} *= \left(\boldsymbol{f_1}^{(\mathbf{1})*} \quad \boldsymbol{f_2}^{(\mathbf{1})*} \quad \dots \quad \boldsymbol{f_m}^{(\mathbf{1}*)}\right)$, didapatkan nilai $\boldsymbol{f}_1^{(1)*}, \boldsymbol{f}_2^{(1)*}, \dots, \boldsymbol{f}_m^{(1)*}$. Selanjutnya berdasarkan $\boldsymbol{f}_1^{(1)*}, \boldsymbol{f}_2^{(1)*}, \dots, \boldsymbol{f}_m^{(1)*}$ akan dicari $\boldsymbol{f}_k^{(1)}$, di mana k = 1, 2, …, m, yang memenuhi $\boldsymbol{f}_k^{(1)^{\boldsymbol{t}}} \boldsymbol{f}_{k'}^{(1)} = 0$ untuk k = k' atau $\boldsymbol{f}_1^{(1)}, \boldsymbol{f}_2^{(1)}, \dots, m$ yang saling ortogonal. Agar memenuhi syarat tersebut, akan digunakan metode *Gram-Schmidt* kepada $\boldsymbol{f}_1^{(1)*}, \boldsymbol{f}_2^{(1)*}, \dots, \boldsymbol{f}_m^{(1)*}$.

Secara umum, nilai $\boldsymbol{a}_j$, $\boldsymbol{v}_j$, $\boldsymbol{q}_j$, dan $\boldsymbol{f_k}$ untuk j = 1, 2, …, p, dan k = 1, 2, …, m, pada iterasi tahap ke-i didapat dengan rumusan sebagai berikut:

1. Cari $\boldsymbol{a_j}^{(i)} = \boldsymbol{F}^{(i-\mathbf{1})^{\boldsymbol{t}}} \boldsymbol{q_j}^{(i-\mathbf{1})} \left(\boldsymbol{q_j}^{(i-\mathbf{1})^{\boldsymbol{t}}} \boldsymbol{q_j}^{(i-\mathbf{1})}\right)^{-\mathbf{1}}$ untuk j = 1, 2, …, p
2. Cari $\boldsymbol{v_j}^{(i)} = \left(\boldsymbol{G_j}^{\boldsymbol{t}} \boldsymbol{G_j}\right)^{-\mathbf{1}} \boldsymbol{G_j}^{\boldsymbol{t}} \boldsymbol{F}^{(i-\mathbf{1})} \boldsymbol{a_j}^{(i)} \left(\boldsymbol{a_j}^{(i)^{\boldsymbol{t}}} \boldsymbol{a_j}^{(i)}\right)^{-\mathbf{1}}$ untuk j = 1, 2, …, p
3. Cari $\boldsymbol{q_j}^{(i)} = \boldsymbol{G_j} \boldsymbol{v_j}^{(i)}$ untuk j = 1, 2, …, p
4. Cari $\boldsymbol{F}^{(i)} *= \left(\boldsymbol{f_1}^{(i)*} \quad \boldsymbol{f_2}^{(i)*} \quad \dots \quad \boldsymbol{f_m}^{(i)*}\right) = \frac{1}{p} \sum_{j=1}^{p} \boldsymbol{q_j}^{(i)} \boldsymbol{a_j}^{(i)^{\boldsymbol{t}}}$. Kemudian lakukan metode *Gram-Schmidt* kepada setiap $\boldsymbol{f}_1^{(1)}, \boldsymbol{f}_2^{(1)}, \dots, \boldsymbol{f}_m^{(1)}$.

Setelah tahap iterasi, selanjutnya adalah menentukan aturan konvergensi (*stopping rule*). Jika hasil yang diperoleh pada tahap iterasi belum memenuhi kriteria pada *stopping rule*, maka tahap iterasi akan diulang kembali hingga hasil yang diperoleh memenuhi kriteria pada *stopping rule*.

Seharusnya, iterasi akan dihentikan jika $\left|\boldsymbol{a_j}^{(i)} - \boldsymbol{a_j}^{(i-\mathbf{1})}\right| < \varepsilon$, $\left|\boldsymbol{q_j}^{(i)} - \boldsymbol{q_j}^{(i-\mathbf{1})}\right| < \varepsilon$, dan $\left|\boldsymbol{f_k}^{(i)} - \boldsymbol{f_k}^{(i-\mathbf{1})}\right| < \varepsilon$, di mana j = 1, 2, …, p, dan k = 1, 2, …, m, atau iterasi akan dihentikan jika $\left|Z_j^{(i)} - Z_j^{(i-\mathbf{1})}\right| < \varepsilon$, di mana $\boldsymbol{\varepsilon}$ adalah suatu bilangan real yang dipilih. Akan dibuktikan bahwa perubahan dari Z_j pada iterasi tahap ke-i dan tahap ke-(i-1) hanya dipengaruhi oleh $\sum_{k=1}^{m} a_{jk}{}^2$ pada iterasi tahap ke-I dan tahap ke-(i-1).

Bukti:

$$Z_j = tr\left[\left(\boldsymbol{q}_j\boldsymbol{\alpha}_j{}^t - \boldsymbol{F}\right)^t\left(\boldsymbol{q}_j\boldsymbol{\alpha}_j{}^t - \boldsymbol{F}\right)\right] \quad (3.2.26)$$

untuk j = 1, 2, …, p. Z_j pada persamaan (3.2.26) dapat ditulis kembali sebagai berikut

$$Z_j = tr\left[\left(\boldsymbol{q}_j\boldsymbol{\alpha}_j{}^t - \boldsymbol{F}\right)^t\left(\boldsymbol{q}_j\boldsymbol{\alpha}_j{}^t - \boldsymbol{F}\right)\right]$$

$$= tr\left[\boldsymbol{\alpha}_j\boldsymbol{q}_j{}^t\boldsymbol{q}_j\boldsymbol{\alpha}_j{}^t - \boldsymbol{\alpha}_j\boldsymbol{q}_j{}^t\boldsymbol{F} - \boldsymbol{F}^t\boldsymbol{q}_j\boldsymbol{\alpha}_j{}^t + \boldsymbol{F}^t\boldsymbol{F}\right]$$

$$= tr\left[\boldsymbol{\alpha}_j\boldsymbol{q}_j{}^t\boldsymbol{q}_j\boldsymbol{\alpha}_j{}^t\right] - tr\left[\boldsymbol{\alpha}_j\boldsymbol{q}_j{}^t\boldsymbol{F}\right] - tr\left[\boldsymbol{F}^t\boldsymbol{q}_j\boldsymbol{\alpha}_j{}^t\right] + tr\left[\boldsymbol{F}^t\boldsymbol{F}\right] \quad (3.2.27)$$

Karena $tr\left[\boldsymbol{\alpha}_j\boldsymbol{q}_j{}^t\boldsymbol{F}\right] = tr\left[\left(\boldsymbol{\alpha}_j\boldsymbol{q}_j{}^t\boldsymbol{F}\right)^t\right] = tr\left[\boldsymbol{F}^t\boldsymbol{q}_j\boldsymbol{\alpha}_j{}^t\right]$ maka

$$= tr\left[\boldsymbol{\alpha}_j\boldsymbol{q}_j{}^t\boldsymbol{q}_j\boldsymbol{\alpha}_j{}^t\right] - tr\left[\boldsymbol{\alpha}_j\boldsymbol{q}_j{}^t\boldsymbol{F}\right] - tr\left[\boldsymbol{F}^t\boldsymbol{q}_j\boldsymbol{\alpha}_j{}^t\right] + tr\left[\boldsymbol{F}^t\boldsymbol{F}\right]$$

$$= tr\left[\boldsymbol{\alpha}_j\boldsymbol{q}_j{}^t\boldsymbol{q}_j\boldsymbol{\alpha}_j{}^t\right] - 2tr\left[\boldsymbol{F}^t\boldsymbol{q}_j\boldsymbol{\alpha}_j{}^t\right] + tr\left[\boldsymbol{F}^t\boldsymbol{F}\right] \quad (3.2.28)$$

Pada subbab sebelumnya, telah diperoleh bahwa nilai $\boldsymbol{\alpha}_j$ yang meminimumkan Z_j adalah

$$\boldsymbol{\alpha}_j = \boldsymbol{F}^t\boldsymbol{q}_j\left(\boldsymbol{q}_j{}^t\boldsymbol{q}_j\right)^{-1} \quad (3.2.29)$$

Sehingga, $\boldsymbol{F}^t\boldsymbol{q}_j = \boldsymbol{\alpha}_j\boldsymbol{q}_j{}^t\boldsymbol{q}_j$. Z_j pada (3.2.29) dapat ditulis kembali sebagai berikut

$$= tr\left[\boldsymbol{\alpha}_j\boldsymbol{q}_j{}^t\boldsymbol{q}_j\boldsymbol{\alpha}_j{}^t\right] - 2tr\left[\left(\boldsymbol{F}^t\boldsymbol{q}_j\right)\boldsymbol{\alpha}_j{}^t\right] + tr\left[\boldsymbol{F}^t\boldsymbol{F}\right]$$

$$= tr\left[\boldsymbol{\alpha}_j\boldsymbol{q}_j{}^t\boldsymbol{q}_j\boldsymbol{\alpha}_j{}^t\right] - 2tr\left[\left(\boldsymbol{\alpha}_j\boldsymbol{q}_j{}^t\boldsymbol{q}_j\right)\boldsymbol{\alpha}_j{}^t\right] + tr\left[\boldsymbol{F}^t\boldsymbol{F}\right]$$

$$= tr\left[\boldsymbol{\alpha}_j\boldsymbol{q}_j{}^t\boldsymbol{q}_j\boldsymbol{\alpha}_j{}^t\right] - 2tr\left[\boldsymbol{\alpha}_j\boldsymbol{q}_j{}^t\boldsymbol{q}_j\boldsymbol{\alpha}_j{}^t\right] + tr\left[\boldsymbol{F}^t\boldsymbol{F}\right]$$

$$= tr\left[\boldsymbol{F}^t\boldsymbol{F}\right] - tr\left[\boldsymbol{\alpha}_j\boldsymbol{q}_j{}^t\boldsymbol{q}_j\boldsymbol{\alpha}_j{}^t\right] \quad (3.2.30)$$

Karena $\boldsymbol{F} = (\boldsymbol{f}_1 \quad \boldsymbol{f}_2 \quad \cdots \quad \boldsymbol{f}_m)$, maka

$$\boldsymbol{F}^t\boldsymbol{F} = \begin{pmatrix} \boldsymbol{f_1}^t \\ \boldsymbol{f_2}^t \\ \vdots \\ \boldsymbol{f_m}^t \end{pmatrix} (\boldsymbol{f_1} \quad \boldsymbol{f_2} \quad \cdots \quad \boldsymbol{f}_m)$$

$$= \begin{pmatrix} \boldsymbol{f_1}^t\boldsymbol{f_1} & \boldsymbol{f_1}^t\boldsymbol{f_2} & \cdots & \boldsymbol{f_1}^t\boldsymbol{f}_m \\ \boldsymbol{f_2}^t\boldsymbol{f_1} & \boldsymbol{f_2}^t\boldsymbol{f_2} & \cdots & \boldsymbol{f_2}^t\boldsymbol{f}_m \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{f}_m{}^t\boldsymbol{f_1} & \boldsymbol{f}_m{}^t\boldsymbol{f_2} & \cdots & \boldsymbol{f}_m{}^t\boldsymbol{f}_m \end{pmatrix} \quad (3.2.31)$$

Pada subbab sebelumnya, telah dijelaskan bahwa $\boldsymbol{f_k}^t\boldsymbol{f_{k'}} = 0$ untuk $k \neq k'$. Jika elemen-elemen dari $\boldsymbol{f_k}$, untuk k = 1, 2, …, m, adalah elemen-elemen yang telah distandarisasi, maka variansi dari F_k, untuk k = 1, 2, …, k adalah

$$var(\,F_k) = \frac{1}{n-1}\boldsymbol{f_k}^t\boldsymbol{f_k} = \frac{1}{n-1}\sum_{i=1}^{n} f_{ik} = 1 \quad (3.2.32)$$

maka $\boldsymbol{f_k}^t\boldsymbol{f_k} = (n-1)$. Jadi, jika elemen-elemen dari $\boldsymbol{f_k}$ adalah elemen-elemen telah distandarisasi, untuk k = 1, 2, …, m, maka $\boldsymbol{F}^t\boldsymbol{F}$ dapat ditulis sebagai berikut:

$$\boldsymbol{F}^t\boldsymbol{F} = \begin{pmatrix} \boldsymbol{f_1}^t\boldsymbol{f_1} & \boldsymbol{f_1}^t\boldsymbol{f_2} & \cdots & \boldsymbol{f_1}^t\boldsymbol{f}_m \\ \boldsymbol{f_2}^t\boldsymbol{f_1} & \boldsymbol{f_2}^t\boldsymbol{f_2} & \cdots & \boldsymbol{f_2}^t\boldsymbol{f}_m \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{f}_m{}^t\boldsymbol{f_1} & \boldsymbol{f}_m{}^t\boldsymbol{f_2} & \cdots & \boldsymbol{f}_m{}^t\boldsymbol{f}_m \end{pmatrix}$$

$$= \begin{pmatrix} (n-1) & 0 & \dots & 0 \\ 0 & (n-1) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & (n-1) \end{pmatrix}$$

$$= (n-1)\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} = (n-1)\boldsymbol{I_m} \quad (3.2.33)$$

di mana $\boldsymbol{I_m}$ adalah matriks identitas berukuran (m x m).

Jika, elemen-elemen dari $\boldsymbol{q_j}$, di mana j = 1, 2, …, p, juga distandarisasi, maka juga berlaku $\boldsymbol{q_j}^t\boldsymbol{q_j} = (n-1)$, untuk j = 1, 2, …, p.

Jika elemen-elemen $\boldsymbol{f_k}$ dan $\boldsymbol{q_j}$, di mana k = 1, 2, …, m, dan j = 1, 2, …, p, adalah elemen-elemen yang telah distandarisasi, maka Z_j pada (3.2.30) dapat ditulis sebagai

$$
\begin{aligned}
Z_j &= \boldsymbol{tr}\left[\left(\boldsymbol{q}_j\boldsymbol{a}_j{}^{\boldsymbol{t}} - \boldsymbol{F}\right)^{\boldsymbol{t}}\left(\boldsymbol{q}_j\boldsymbol{a}_j{}^{\boldsymbol{t}} - \boldsymbol{F}\right)\right] \\
&= \boldsymbol{tr}[\boldsymbol{F}^{\boldsymbol{t}}\boldsymbol{F}] - \boldsymbol{tr}\left[\boldsymbol{a}_j\boldsymbol{q}_j{}^{\boldsymbol{t}}\boldsymbol{q}_j\boldsymbol{a}_j{}^{\boldsymbol{t}}\right] \\
&= tr[\boldsymbol{F}^{\boldsymbol{t}}\boldsymbol{F}] - tr\left[\boldsymbol{a}_j\left(\boldsymbol{q}_j{}^{\boldsymbol{t}}\boldsymbol{q}_j\right)\boldsymbol{a}_j{}^{\boldsymbol{t}}\right] \\
&= \boldsymbol{tr}[(n-1)\boldsymbol{I}_m] - \boldsymbol{tr}\left[\boldsymbol{a}_j(n-1)\boldsymbol{a}_j{}^{\boldsymbol{t}}\right] \\
&= (n-1)m - (n-1)\boldsymbol{tr}\left[\boldsymbol{a}_j\boldsymbol{a}_j{}^{\boldsymbol{t}}\right]
\end{aligned}
\tag{3.2.34}
$$

Karena $\boldsymbol{a}_j = \begin{pmatrix} a_{j1} \\ a_{j2} \\ \vdots \\ a_{jm} \end{pmatrix}$, untuk setiap j = 1, 2, …, p, maka

$$
\begin{aligned}
\boldsymbol{a}_j\boldsymbol{a}_j{}^{\boldsymbol{t}} &= \begin{pmatrix} a_{j1} \\ a_{j2} \\ \vdots \\ a_{jm} \end{pmatrix}\begin{pmatrix} a_{j1} & a_{j2} & \cdots & a_{jm} \end{pmatrix} \\
&= \begin{pmatrix} a_{j1}{}^{(2)} & a_{j1}a_{j2} & \cdots & a_{j1}a_{jm} \\ a_{j2}a_{j1} & a_{j2}{}^{(2)} & \cdots & a_{j2}a_{jm} \\ \vdots & \vdots & \ddots & \vdots \\ a_{jm}a_{j1} & a_{jm}a_{j2} & \cdots & a_{jm}{}^{(2)} \end{pmatrix}
\end{aligned}
\tag{3.2.35}
$$

Sehingga, $\boldsymbol{tr}\left[\boldsymbol{a}_j\boldsymbol{a}_j{}^{\boldsymbol{t}}\right]$ adalah

$$
\boldsymbol{tr}\left[\boldsymbol{a}_j\boldsymbol{a}_j{}^{\boldsymbol{t}}\right] = \boldsymbol{tr}\begin{pmatrix} a_{j1}{}^{(2)} & a_{j1}a_{j2} & \cdots & a_{j1}a_{jm} \\ a_{j2}a_{j1} & a_{j2}{}^{(2)} & \cdots & a_{j2}a_{jm} \\ \vdots & \vdots & \ddots & \vdots \\ a_{jm}a_{j1} & a_{jm}a_{j2} & \cdots & a_{jm}{}^{(2)} \end{pmatrix} = \textstyle\sum_{k=1}^{m} a_{jk}{}^{2}
\tag{3.2.36}
$$

Sehingga, Z_j pada (3.2.34) dapat ditulis sebagai

$$
Z_j = \boldsymbol{tr}\left[\left(\boldsymbol{q}_j\boldsymbol{a}_j{}^{\boldsymbol{t}} - \boldsymbol{F}\right)^{\boldsymbol{t}}\left(\boldsymbol{q}_j\boldsymbol{a}_j{}^{\boldsymbol{t}} - \boldsymbol{F}\right)\right]
$$

$$= (n-1)m - (n-1)tr\left[\boldsymbol{a}_j \boldsymbol{a}_j{}^{\boldsymbol{t}}\right]$$

$$= (n-1)m - (n-1)\sum_{k=1}^{m} a_{jk}{}^2$$

$$= (n-1)\left(m - \sum_{k=1}^{m} a_{jk}{}^2\right) \tag{3.2.37}$$

Berdasarkan (3.2.37), jika elemen-elemen $\boldsymbol{f_k}$ dan $\boldsymbol{q_j}$, di mana k = 1, 2, …, m, dan j = 1, 2, …, p, adalah elemen-elemen yang telah distandarisasi, maka perubahan dari Z_j, untuk j = 1, 2, …, p, pada iterasi tahap ke-i dan tahap ke-(i-1) hanya ditentukan oleh $\sum_{k=1}^{m} a_{jk}{}^2$ pada tahap ke-i dan tahap ke-(i-1). Oleh karena itu, aturan konvergensi (*stopping rule*) hanya ditentukan oleh perubahan $\sum_{k=1}^{m} a_{jk}{}^2$ pada tahap ke-i dan tahap ke-(i-1), untuk j = 1, 2, …, p.

Stopping rule untuk iterasi tahap ke–i ditentukan sebagai berikut: Jika $\sum_{j=1}^{p}\sum_{k=1}^{m} a_{jk}{}^{(i)^2}$ pada tahap ke-i memenuhi

$$\sum_{j=1}^{p}\sum_{k=1}^{m} a_{jk}{}^{(i-1)^2} - \sum_{j=1}^{p}\sum_{k=1}^{m} a_{jk}{}^{(i)^2} < \varepsilon \tag{3.2.38}$$

di mana ε merupakan suatu nilai tertentu yang dipilih, maka proses iterasi dihentikan. Tetapi, jika belum memenuhi, atau jika $\sum_{j=1}^{p}\sum_{k=1}^{m} a_{jk}{}^{(i-1)^2} - \sum_{j=1}^{p}\sum_{k=1}^{m} a_{jk}{}^{(i)^2} \geq \varepsilon$, maka proses dilanjutkan ke iterasi tahap berikutnya.

Selanjutnya, akan dibuktikan bahwa $\sum_{j=1}^{p}\sum_{k=1}^{m} a_{jk}{}^2$ menyatakan total variansi dari $Q_1, Q_2, \dots, Q_m$ yang dicakup oleh komponen-komponen utama $F_1, F_2, \dots, F_m$.

Bukti:

$$\boldsymbol{q_j} = a_{j1}\boldsymbol{f_1} + a_{j2}\boldsymbol{f_2} + \cdots + a_{jm}\,\boldsymbol{f}_m = \boldsymbol{F}\boldsymbol{a}_j \tag{3.2.39}$$

untuk j = 1, 2, …, p.

Sebut, $\boldsymbol{Q} = (\boldsymbol{q_1} \quad \boldsymbol{q_2} \quad \cdots \quad \boldsymbol{q}_p)$, dan $\boldsymbol{A} = (\boldsymbol{a_1} \quad \boldsymbol{a_2} \quad \cdots \quad \boldsymbol{a}_p)$,

$$\boldsymbol{Q} = (\boldsymbol{q_1} \quad \boldsymbol{q_2} \quad \cdots \quad \boldsymbol{q}_p) = (\boldsymbol{F}\boldsymbol{a_1} \quad \boldsymbol{F}\boldsymbol{a_2} \quad \cdots \quad \boldsymbol{F}\boldsymbol{a}_p),$$

$$= \boldsymbol{F}(\boldsymbol{a_1} \quad \boldsymbol{a_2} \quad \cdots \quad \boldsymbol{a}_p) = \boldsymbol{F}\boldsymbol{A} \tag{3.2.40}$$

Selanjutnya, akan dicari $\boldsymbol{Q^tQ}$.

$$\boldsymbol{Q^tQ} = (\boldsymbol{FA})^t\boldsymbol{FA} = \boldsymbol{A^tF^tFA} \tag{3.2.41}$$

Telah dijelaskan sebelumnya bahwa elemen-elemen dari $\boldsymbol{f_1}, \boldsymbol{f_2}, \dots, \boldsymbol{f_m}$, dan $\boldsymbol{q_1}, \boldsymbol{q_2}, \dots, \boldsymbol{q_p}$ telah distandarisasi. Berdasarkan (3.2.33), $\boldsymbol{F^tF} = (n-1)\boldsymbol{I_m}$, maka (3.2.41) dapat ditulis sabagai

$$\boldsymbol{Q^tQ} = \boldsymbol{A^tF^tFA} = \boldsymbol{A^t}(n-1)\boldsymbol{I_m}\,\boldsymbol{A} = (n-1)\boldsymbol{A^tA} \tag{3.2.44}$$

Misalkan, $\boldsymbol{\Sigma}$ adalah matriks variansi-kovariansi dari $Q_1, Q_2, \dots, Q_m$. Matriks $\boldsymbol{\Sigma}$ dicari dengan $\boldsymbol{\Sigma} = \frac{1}{n-1}\boldsymbol{Q^tQ}$

$$\boldsymbol{\Sigma} = \frac{1}{n-1}\boldsymbol{Q^tQ} = \frac{1}{n-1}(n-1)\boldsymbol{A^tA} = \boldsymbol{A^tA} \tag{3.2.45}$$

Total variansi dari $Q_1, Q_2, \dots, Q_m$ diperoleh dengan $\boldsymbol{tr}(\boldsymbol{\Sigma}) = \boldsymbol{tr}(\boldsymbol{A^tA})$. $\boldsymbol{A^tA}$ dapat ditulis sebagai berikut

$$\boldsymbol{A^tA} = \begin{pmatrix} \boldsymbol{a_1}^t \\ \boldsymbol{a_2}^t \\ \vdots \\ \boldsymbol{a_p}^t \end{pmatrix} (\boldsymbol{a_1} \quad \boldsymbol{a_2} \quad \cdots \quad \boldsymbol{a_p})$$

$$= \begin{pmatrix} \boldsymbol{a_1}^t\boldsymbol{a_1} & \boldsymbol{a_1}^t\boldsymbol{a_2} & \cdots & \boldsymbol{a_1}^t\boldsymbol{a_p} \\ \boldsymbol{a_2}^t\boldsymbol{a_1} & \boldsymbol{a_2}^t\boldsymbol{a_2} & \cdots & \boldsymbol{a_2}^t\boldsymbol{a_p} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{a_p}^t\boldsymbol{a_1} & \boldsymbol{a_p}^t\boldsymbol{a_2} & \dots & \boldsymbol{a_p}^t\boldsymbol{a_p} \end{pmatrix} \tag{3.2.46}$$

Berdasarkan 3.2.46, $\boldsymbol{tr}(\boldsymbol{\Sigma}) = \boldsymbol{tr}(\boldsymbol{A^tA})$ diperoleh dengan

$$\boldsymbol{tr}(\boldsymbol{\Sigma}) = \boldsymbol{tr}(\boldsymbol{A^tA}) = \boldsymbol{tr}\begin{pmatrix} \boldsymbol{a_1}^t\boldsymbol{a_1} & \boldsymbol{a_1}^t\boldsymbol{a_2} & \cdots & \boldsymbol{a_1}^t\boldsymbol{a_p} \\ \boldsymbol{a_2}^t\boldsymbol{a_1} & \boldsymbol{a_2}^t\boldsymbol{a_2} & \cdots & \boldsymbol{a_2}^t\boldsymbol{a_p} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{a_p}^t\boldsymbol{a_1} & \boldsymbol{a_p}^t\boldsymbol{a_2} & \dots & \boldsymbol{a_p}^t\boldsymbol{a_p} \end{pmatrix}$$

$$= \boldsymbol{a_1}^t\boldsymbol{a_1} + \boldsymbol{a_2}^t\boldsymbol{a_2} + \cdots + \boldsymbol{a_p}^t\boldsymbol{a_p}$$

$$= \sum_{k=1}^{m} a_{1k}{}^2 + \sum_{k=1}^{m} a_{1k}{}^2 + \cdots + \sum_{k=1}^{m} a_{1k}{}^2$$

$$= \sum_{j=1}^{p} \sum_{k=1}^{m} a_{jk}^{2} \qquad (3.2.47)$$

Terbukti bahwa $\sum_{j=1}^{p} \sum_{k=1}^{m} a_{jk}^{2}$ menyatakan total variansi dari $Q_1, Q_2, \dots, Q_m$ yang dicakup dalam komponen-komponen utama $F_1, F_2, \dots, F_m$.

Pada subbab sebelumnya dijelaskan bahwa $\boldsymbol{f_k}$, $\boldsymbol{a_j}$ dan $\boldsymbol{q_j}$ di mana k = 1, 2, …, m, dan j = 1, 2, …, p, dicari dengan meminimumkan

$$tr\left[\left(\boldsymbol{q}_j\boldsymbol{a}_j{}^t - \boldsymbol{F}\right)^t\left(\boldsymbol{q}_j\boldsymbol{a}_j{}^t - \boldsymbol{F}\right)\right] \qquad (3.2.48)$$

Jika elemen-elemen dari $\boldsymbol{f_k}$ dan $\boldsymbol{q_j}$ adalah elemen-elemen yang distandarisasi, maka telah dibuktikan sebelumnya pada (3.2.37) bahwa

$$\boldsymbol{tr}\left[\left(\boldsymbol{q}_j\boldsymbol{a}_j{}^t - \boldsymbol{F}\right)^t\left(\boldsymbol{q}_j\boldsymbol{a}_j{}^t - \boldsymbol{F}\right)\right] = (n-1)\left(m - \sum_{k=1}^{m} a_{jk}^{2}\right) \qquad (3.2.49)$$

Berarti, meminimumkan $\boldsymbol{tr}\left[\left(\boldsymbol{q}_j\boldsymbol{a}_j{}^t - \boldsymbol{F}\right)^t\left(\boldsymbol{q}_j\boldsymbol{a}_j{}^t - \boldsymbol{F}\right)\right]$ sama seperti meminimumkan $(n-1)\left(m - \sum_{k=1}^{m} a_{jk}^{2}\right)$, di mana elemen-elemen dari $\boldsymbol{f_k}$ dan $\boldsymbol{q_j}$ adalah elemen-elemen yang distandarisasi. Karena n dan m adalah suatu nilai yang konstan, maka meminimumkan $(n-1)\left(m - \sum_{k=1}^{m} a_{jk}^{2}\right)$ sama seperti memaksimumkan $\sum_{k=1}^{m} a_{jk}^{2}$ untuk j = 1, 2, …, p.

Karena $\sum_{j=1}^{p} \sum_{k=1}^{m} a_{jk}^{2}$ menyatakan total variansi dari $Q_1, Q_2, \dots, Q_m$ yang dicakup dalam komponen-komponen utama $F_1, F_2, \dots, F_m$, maka dapat disimpulkan bahwa tahap iterasi pada metode PRINCALS akan memaksimumkan total variansi dari $Q_1, Q_2, \dots, Q_m$ yang dicakup dalam komponen-komponen utama $F_1, F_2, \dots, F_m$.

Karena elemen-elemen dari $\boldsymbol{q_j}$ distandarisasi, maka variansi dari Q_j untuk j = 1, 2, …, p adalah satu, sehingga total variansi dari $Q_1, Q_2, \dots, Q_m$ adalah p. Persentase total variansi dari $Q_1, Q_2, \dots, Q_m$ yang dicakup oleh $F_1, F_2, \dots, F_m$ dicari dengan

$$\frac{\sum_{j=1}^{p}\sum_{k=1}^{m} a_{jk}^{2}}{p} \; x \; 100\% \qquad (3.2.50)$$

Selanjutnya, akan dibuktikan bahwa koefisien a_{jk} untuk variabel ke-j, yaitu Q_j, pada komponen utama ke-k, yaitu F_k, di mana k = 1, 2, …, m, dan j= 1, 2, …, p menyatakan korelasi antara Q_j dan F_k.

Bukti:

$$Q_j = a_{j1}F_1 + a_{j2}F_2 + \cdots + a_{jm}\,F_m \qquad (3.2.51)$$

$$corr(Q_j, F_k) = corr(a_{j1}F_1 + a_{j2}F_2 + \cdots + a_{jm}\,F_m, F_k)$$

$$= corr(a_{j1}F_1, F_k) + corr(a_{j2}F_2, F_k) + \cdots + corr(a_{jk}F_k, F_k) + \cdots + corr(a_{jm}F_m, F_k)$$

$$= a_{j1}corr(F_1, F_k) + a_{j2}corr(F_2, F_k) + \cdots + a_{jk}corr(F_k, F_k) + \cdots + a_{jm}corr(F_m, F_k)$$

$$= a_{jk}corr(F_k, F_k) = a_{jk} \qquad (3.2.52)$$

Terbukti bahwa koefisien a_{jk} menyatakan korelasi antara antara Q_j dan F_k, di mana k = 1, 2, …, m, dan j= 1.

Untuk merepresentasikan komponen utama $F_1, F_2, \ldots, F_m$, dilakukan dengan melihat korelasi antara Q_j dan F_k, di mana k = 1, 2, …, m, dan j= 1, 2, …, p, atau melihat pada koefisien a_{jk}, di mana telah dibuktikan bahwa a_{jk} menyatakan korelasi antara Q_j dan F_k, di mana k = 1, 2, …, m, dan j= 1. Jika korelasi antara Q_j dan F_k, atau a_{jk} besar, maka komponen utama F_k dikatakan menekankan pada variabel Q_j, di mana k = 1, 2, …, m, dan j= 1.

3.2 Pembentukan Indikator Sasaran dengan *Proxy Means Test* Berdasarkan Metode PRINCALS

Indikator sasaran yang terbentuk dengan *proxy means test* harus merupakan variabel numerik agar indikator sasaran tersebut dapat memberikan informasi yang lebih banyak pada analisis selanjutnya dibandingkan variabel kategorik. Telah dijelaskan pada bab sebelumnya bahwa pada saat variabel-variabel proxi merupakan variabel kategorik, maka salah satu metode yang digunakan untuk membentuk indikator sasaran dengan *proxy means test* adalah metode PRINCALS yang telah dijelaskan pada subbab sebelumnya. Jika metode PRINCALS digunakan dalam *proxy means test*, maka indikator sasaran dipilih dari satu atau beberapa komponen utama yang didapat sesuai dengan kebutuhan peneliti.

BAB 4

CONTOH PENERAPAN

Pada bab ini akan dibahas contoh mengenai pembentukkan indikator sasaran dengan *proxy means test* berdasarkan metode PRINCALS.

Sampai saat ini, penghasilan suatu keluarga masih dianggap sebagai salah satu kriteria penentu apakah suatu keluarga dikategorikan miskin atau tidak. Hal ini terlihat dari kriteria BPS dalam menentukan tingkat kemiskinan di mana penghasilan keluarga termasuk dalam kriteria tersebut. Secara subtantif, semakin kecil penghasilan suatu keluarga, maka semakin mungkin keluarga tersebut masuk ke dalam kategori miskin. Namun, pada saat pengambilan sampel, data mengenai penghasilan keluarga cenderung mudah dimanipulasi. Terutama pengambilan data untuk penelitian sosial, seperti pemberian bantuan pemerintah. Responden cenderung tidak memberikan secara benar data mengenai penghasilan mereka. Jika penentuan tingkat kemiskinan masih menggunakan data penghasilan keluarga yang termanipulasi, maka hasil yang didapat kurang baik.

Berdasarkan fenomena di atas penulis mencoba untuk menentukan indikator sasaran dari variabel tingkat kemiskinan dengan *proxy means test* berdasarkan metode PRINCALS. Selain mendapatkan indikator sasaran (yang tidak termanipulasi) untuk mengukur tingkat kemiskinan, akan dilihat juga korelasi antara indikator sasaran dengan variabel total income rumah tangga. Jika tidak ada korelasi antara indikator sasaran dengan variabel total income rumah tangga, maka dapat disimpulkan bahwa total income rumah tangga termanipulasi. Pengujian korelasi antara indikator sasaran dengan total income rumah tangga dilakukan dengan Spearmen-Rho dengan tingkat signifikansi 0.150.

Seperti telah dijelaskan sebelumnya bahwa dalam membangun indikator sasaran, pertama kali harus ditentukan sejumlah variabel proxi, di mana variabel proxi tersebut secara subtantif berhubungan erat dengan variabel sasaran (kemiskinan), mudah diukur dan tidak dapat dimanipulasi. Berdasarkan penelitian yang dilakukan oleh suatu lembaga yang memperhatikan penanggulangan kemiskinan di Indonesia, karakteristik tempat tinggal suatu keluarga, yaitu jenis lantai terluas, jenis dinding terluas, jenis atap terluas, penggunaan jamban/kakus, sumber penerangan utama, dan bahan bakar utama merupakan kriteria yang dapat menentukan tingkat kemiskinan. Untuk itu, penulis memilih keenam variabel tersebut sebagai variabel proxi, di mana keenam variabel tersebut merupakan variabel kategorik karena variabel-variabel tersebut memenuhi syarat untuk menjadi variabel proxi, yaitu mudah diukur, dan tidak mudah dimanipulasi pada saat pengukuran.

4.1 Data

Data yang digunakan adalah data sekunder dari penelitian suatu lembaga yang memperhatikan penanggulangan kemiskinan di Indonesia. Data tersebut mencakup empat provinsi, yaitu Kalimantan Barat, Nusa Tenggara Bata, Sulawesi Selatan, dan Maluku. Analisis akan dilakukan untuk masing-masing provinsi. Jumlah data yang digunakan untuk masing-masing provinsi adalah sebagai berikut:

- ✓ Kalimantan Barat = 163 rumah tangga
- ✓ Nusa Tenggara Barat = 168 rumah tangga
- ✓ Sulawesi Selatan = 169 rumah tangga
- ✓ Maluku = 167 rumah tangga

Variabel-variabel proxi yang digunakan adalah sebagai berikut:

- Jenis lantai terluas : 1, bambu/tanah
  2, keramik/semen/kayu jelek
  3, keramik/semen/kayu bagus

- Jenis dinding terluas : 1, bambu/rumbia

2, tembok/kayu kondisi jelek

3, tembok/kayu kondisi baik

- Jenis atap : 1, rumbia/ilalang/jerami

  2, sirap/genteng/seng/asbes kondisi jelek

  3, sirap/genteng/seng/asbes kondisi baik

- Penggunaan jamban : 1, bersama/umum

  2, sendiri

- Sumber penerangan utama : 1, bukan listrik

  2, listrik tanpa meteran

  3, listrik dengan meteran

- Bahan bakar utama : 1, kayu/arang

  2, minyak tanah

  3, gas/listrik

- Total income rumah tangga dalam seminggu

Nilai numerik pada setiap kategori untuk setiap variabel akan dijadikan nilai numerik awal untuk metode PRINCALS. Persentase total variansi dari variabel-variabel proxi yang dicakup dalam komponen ditentukan sebesar lebih dari 60%.

4.2 Analisis Data

Analisis data dilakukan untuk masing-masing provinsi dan di dapat hasil sebagai berikut:

4.2.1 Kalimantan Barat

Untuk provinsi Kalimantan Barat, dengan menggunakan 163 rumah tangga, diperoleh hasil sebagai berikut:

1. Diperoleh dua komponen utama yang mencakup 62.2% total variansi dari variabel-variabel proxi.
2. Komponen pertama (F_1) mencakup 34.5% total variansi dari variabel-variabel proxi, sedangkan komponen kedua (F_2) mencakup 27.7%.
3. Kuantifikasi optimal pada setiap kategori untuk setiap variabel adalah sebagai berikut:

Tabel 4.1 Kuantifikasi optimal untuk variabel-variabel proxi provinsi Kalimantan Barat

Variabel	Kategori	Kuantifikasi optimal
Jenis lantai terluas	Bambu/tanah	-4.08
	Keramik/semen/kayu jelek	-0.27
	Keramik/semen/kayu bagus	3
Jenis dinding terluas	Bambu/rumbia	-2.11
	Tembok/kayu kondisi jelek	-0.22
	Tembok/kayu kondisi bagus	3.09
Jenis atap	Rumbia/ilalang/jerami	-1.13
	Sirap/genteng/seng/asbes kondisi jelek	0.52
	Sirap/genteng/seng/asbes kondisi bagus	3.91
Penggunaan jamban/kakus	Bersama/umum	-1.37
	Sendiri	0.73
Sumber penerangan utama	Bukan listrik	-1.93
	Listrik tanpa meteran	-1.71
	Listrik dengan meteran	0.59

Bahan bakar utama	Kayu/arang	1.09
	Minyak tanah	-0.91
	Gas/listrik	-1.5

4. Didapat korelasi antara dua komponen dengan variabel-variabel proxi sebagai berikut:

Tabel 4.2 Korelasi antara F_1 (komponen pertama) dan F_2 (komponen kedua) dengan variabel-variabel proxi untuk provinsi Kalimantan Barat

Variabel	F_1	F_2
Jenis lantai terluas	0.500	-0.585
Jenis dinding terluas	0.506	-0.672
Jenis atap	0.685	-0.245
Penggunaan jamban	0.621	0.450
Sumber penerangan	0.721	0.284
Bahan bakar	0.436	0.724

Berdasarkan tabel di atas, komponen pertama berkorelasi cukup besar dengan variabel proxi jenis atap (0.685), penggunaan jamban/kakus (0.621), dan sumber penerangan utama (0.721). Sedangkan komponen utama kedua berkorelasi cukup besar dengan variabel proxi jenis lantai terluas (-0.585), jenis dinding terluas (-0.672), dan bahan bakar utama (0.724).

Oleh karena itu, tingkat kemiskinan di provinsi Kalimantan Barat dapat diukur dengan indikator sasaran:

- ✓ F_1 yang lebih menekankan pada variabel jenis atap, penggunaan jamban/kakus, dan sumber penerangan utama

dan atau

- ✓ F_2 yang lebih menekankan pada variabel jenis lantai terluas, jenis dinding terluas, dan bahan bakar utama.

F_1 dan F_2 diperoleh dengan formula PRINCALS, yaitu sebagai berikut:

F_1 = (1/6)(0.5jenis_lantai + 0.506jenis_dinding + 0.685jenis_atap + 0.621penggunaan_jamban + 0.721penerangan + 0.436bahan_bakar)

F_2 = (1/6)(-0.585jenis_lantai - 0.672jenis_dinding - 0.245jenis_atap + 0.450penggunaan_jamban + 0.284penerangan + 0.724bahan_bakar)

Untuk analisis selanjutnya, pengkategorian kemiskinan untuk provinsi Kalimantan Barat dapat dilakukan berdasarkan indikator sasaran F_1 dan atau F_2.

5. Pengujian korelasi antara indikator sasaran (F_1 dan F_2) tingkat kemiskinan untuk provinsi Kalimantan Barat dengan total income rumah tangga

Tabel 4.3 Output uji korelasi Spearman's rho antara indikator sasaran tingkat kemiskinan dengan total income rumah tangga provinsi Kalimantan Barat

Correlations

			Total income rumah tangga rata-rata dalam seminggu	Komponen Kalimantan Barat 1	Komponen Kalimantan Barat 2
Spearman's rho	Total income rumah tangga rata-rata dalam seminggu	Correlation Coefficient	1.000	.158	-.263
		Sig. (2-tailed)	.	.044	.001
		N	669	163	163
	Komponen Kalimantan Barat 1	Correlation Coefficient	.158	1.000	.002
		Sig. (2-tailed)	.044	.	.976
		N	163	163	163
	Komponen Kalimantan Barat 2	Correlation Coefficient	-.263	.002	1.000
		Sig. (2-tailed)	.001	.976	.
		N	163	163	163

Berdasarkan output di atas, nilai $\hat{\alpha}$ antara F_1 dengan total income rumah tangga adalah 0.044, dan antara F_2 dengan total income rumah tangga adalah 0.001, sehingga dengan memilih tingkat signifikansi 0.150, maka dapat disimpulkan bahwa terdapat korelasi antara indikator sasaran (F_1 dan F_2) dengan total income rumah tangga atau total income rumah tangga untuk provinsi

Kalimantan Barat tidak termanipulasi sehingga total income rumah tangga bisa digunakan untuk mengukur tingkat kemiskinan di provinsi Kalimantan Barat.

4.2.2 Nusa Tenggara Barat

Untuk provinsi Nusa Tenggara Barat, dengan menggunakan 168 rumah tangga diperoleh hasil sebagai berikut

1. Diperoleh dua komponen utama yang mencakup 62.71% total variansi dari variabel-variabel proxi.
2. Komponen pertama (F_1) mencakup 44.4% total variansi dari variabel-variabel proxi, sedangkan komponen kedua (F_2) mencakup 18.31%.
3. Kuantifikasi optimal pada setiap kategori untuk setiap variabel adalah sebagai berikut:

Tabel 4.4 Kuantifikasi optimal untuk variabel-variabel proxi provinsi Nusa Tenggara Barat

Variabel	Kategori	Kuantifikasi optimal
Jenis lantai terluas	Bambu/tanah	-1.25
	Keramik/semen/kayu jelek	0.67
	Keramik/semen/kayu bagus	2.84
Jenis dinding terluas	Bambu/rumbia	-1.29
	Tembok/kayu kondisi jelek	0.65
	Tembok/kayu kondisi bagus	3.56
Jenis atap	Rumbia/ilalang/jerami	-1.97
	Sirap/genteng/seng/asbes kondisi jelek	0.39
	Sirap/genteng/seng/asbes kondisi bagus	2.01
Penggunaan jamban/kakus	Bersama/umum	-0.65
	Sendiri	1.54
Sumber penerangan utama	Bukan listrik	-3.25
	Listrik tanpa meteran	0.00
	Listrik dengan meteran	1.86

Bahan bakar utama	Kayu/arang	-0.52
	Minyak tanah	1.92
	Gas/listrik	0.00

4. Didapat korelasi antara dua komponen dengan variabel-variabel proxi adalah sebagai berikut:

Tabel 4.5 Korelasi antara F_1 (komponen pertama) dan F_2 (komponen kedua) dengan variabel-variabel proxi untuk provinsi Nusa Tenggara Barat

Variabel	F_1	F_2
Jenis lantai terluas	-0.876	-0.257
Jenis dinding terluas	-0.867	-0.241
Jenis atap	-0.833	-0.171
Penggunaan jamban	-0.368	0.604
Sumber penerangan	-0.502	0.376
Bahan bakar	-0.251	0.662

Berdasarkan output di atas, komponen pertama berkorelasi cukup besar dengan variabel proxi jenis lantai terluas (-0.876), jenis dinding terluas (0.867), jenis atap (-0.833), dan sumber penerangan utama (-0.502). Sedangkan komponen kedua berkorelasi cukup besar dengan variabel proxi penggunaan jamban (0.604), dan bahan bakar utama (0.662).

Oleh karena itu, tingkat kemiskinan di provinsi Nusa Tenggara Barat dapat diukur dengan indikator sasaran:

- ✓ F_1 yang lebih menekankan pada variabel jenis lantai terluas, jenis dinding terluas, jenis atap, dan sumber penerangan utama

dan atau

- ✓ F_2 yang lebih menekankan pada variabel penggunaan jamban dan bahan bakar utama

F_1 dan F_2 diperoleh dengan formula PRINCALS, yaitu sebagai berikut:

F_1 = (1/6)(-0.876jenis_lantai - 0.867jenis_dinding - 0.833jenis_atap - 0.368penggunaan_jamban - 0.502penerangan - 0.251bahan_bakar)

F_2 = (1/6)(-0.257jenis_lantai - 0.241jenis_dinding - 0.171jenis_atap + 0.604penggunaan_jamban + 0.376penerangan + 0.662bahan_bakar

Untuk analisis selanjutnya, pengkategorian kemiskinan provinsi Nusa Tenggara Barat dapat dilakukan berdasarkan indikator sasaran F_1 dan atau F_2.

5. Pengujian korelasi antara indikator sasaran (F_1 dan F_2) tingkat kemiskinan untuk provinsi Nusa Tenggara Barat dengan total income rumah tangga

Tabel 4.6 Output uji korelasi Spearman's rho antara indikator sasaran tingkat kemiskinan dengan total income rumah tangga provinsi Nusa Tenggara Barat

Correlations

			Total income rumah tangga rata-rata dalam seminggu	Komponen NTB 1	Komponen NTB 2
Spearman's rho	Total income rumah tangga rata-rata dalam seminggu	Correlation Coefficient	1.000	-.062	-.057
		Sig. (2-tailed)	.	.427	.444
		N	669	168	168
	Komponen NTB 1	Correlation Coefficient	-.062	1.000	-.054
		Sig. (2-tailed)	.427	.	.490
		N	168	168	168
	Komponen NTB 2	Correlation Coefficient	-.057	-.054	1.000
		Sig. (2-tailed)	.444	.490	.
		N	168	168	168

Berdasarkan output di atas, nilai $\hat{\alpha}$ antara F_1 dengan total income rumah tangga adalah 0.427, dan antara F_2 dengan total income rumah tangga adalah 0.444, sehingga dengan memilih tingkat signifikansi 0.150, maka dapat disimpulkan bahwa tidak terdapat korelasi antara indikator sasaran (F_1 dan F_2) dengan total income rumah tangga atau total income rumah tangga untuk provinsi Nusa Tenggara Barat termanipulasi sehingga total income rumah tangga tidak bisa digunakan untuk mengukur tingkat kemiskinan di provinsi Nusa Tenggara Barat.

4.2.3 Sulawesi Selatan

Untuk provinsi Sulawesi Selatan, dengan menggunakan 169 rumah tangga, diperoleh hasil sebagai berikut

1. Diperoleh dua komponen utama yang mencakup 67.4% total variansi dari variabel-variabel proxi.
2. Komponen pertama (F_1) mencakup 45.8% total variansi dari variabel-variabel proxi, sedangkan komponen kedua (F_2) mencakup 21.6%.
3. Kuantifikasi optimal pada setiap kategori untuk setiap variabel adalah sebagai berikut:

Tabel 4.7 Kuantifikasi optimal untuk variabel-variabel proxi provinsi Sulawesi Selatan

Variabel	Kategori	Kuantifikasi optimal
Jenis lantai terluas	Bambu/tanah	-0.5
	Keramik/semen/kayu jelek	-0.28
	Keramik/semen/kayu bagus	3.08
Jenis dinding terluas	Bambu/rumbia	-0.49
	Tembok/kayu kondisi jelek	-0.27
	Tembok/kayu kondisi bagus	3.44
Jenis atap	Rumbia/ilalang/jerami	-0.54
	Sirap/genteng/seng/asbes kondisi jelek	-0.2
	Sirap/genteng/seng/asbes kondisi bagus	4.47
Penggunaan jamban/kakus	Bersama/umum	-1.13
	Sendiri	0.89
Sumber penerangan utama	Bukan listrik	-2.42
	Listrik tanpa meteran	-0.63
	Listrik dengan meteran	1.09
Bahan bakar utama	Kayu/arang	-1.03
	Minyak tanah	-0.55
	Gas/listrik	1.09

4. Didapat korelasi antara dua komponen dengan variabel-variabel proxi adalah sebagai berikut:

Tabel 4.8 Korelasi antara F_1 (komponen pertama) dan F_2 (komponen kedua) dengan variabel-variabel proxi untuk provinsi Sulawesi Barat

Variabel	F_1	F_2
Jenis lantai terluas	0.836	-0.259
Jenis dinding terluas	0.852	-0.318
Jenis atap	0.860	-0.299
Penggunaan jamban	0.373	0.685
Sumber penerangan	0.441	0.613
Bahan bakar	0.499	0.439

Berdasarkan output di atas, komponen pertama berkorelasi cukup besar dengan variabel proxi jenis lantai terluas (0.836), jenis dinding terluas (0.852), jenis atap (0.860), dan bahan bakar utama (0.499). Sedangkan komponen kedua berkorelasi cukup besar dengan variabel proxi penggunaan jamban (0.685), dan sumber penerangan utama (0.613).

Oleh karena itu, tingkat kemiskinan di provinsi Sulawesi Selatan dapat diukur dengan indikator sasaran:

- ✓ F_1 yang lebih menekankan pada variabel jenis lantai terluas, jenis dinding terluas, jenis atap, dan bahan bakar utama

dan atau

- ✓ F_2 yang lebih menekankan pada variabel penggunaan jamban dan sumber penerangan utama

F_1 dan F_2 diperoleh dengan formula PRINCALS, yaitu sebagai berikut:

F_1 = (1/6)(0.836jenis_lantai + 0.852jenis_dinding + 0.860jenis_atap + 0.373penggunaan_jamban + 0.441penerangan + 0.499bahan_bakar)

$F_2 =$ (1/6)(-0.259jenis_lantai - 0.318jenis_dinding - 0.299jenis_atap + 0.685penggunaan_jamban + 0.613penerangan + 0.439bahan_bakar)

Untuk analisis selanjutnya, pengkategorian kemiskinan provinsi Sulawesi Selatan dapat dilakukan berdasarkan indikator sasaran F_1 dan atau F_2.

5. Diperoleh korelasi antara indikator sasaran (F_1 dan F_2) tingkat kemiskinan untuk provinsi Sulawesi Selatan dengan total income rumah tangga

Tabel 4.9 Output uji korelasi Spearman's rho antara indikator sasaran tingkat kemiskinan dengan total income rumah tangga provinsi Sulawesi Selatan

Correlations

			Total income rumah tangga rata-rata dalam seminggu	Komponen Sulawesi Selatan 1	Komponen Sulawesi Selatan 2
Spearman's rho	Total income rumah tangga rata-rata dalam seminggu	Correlation Coefficient	1.000	.258	.143
		Sig. (2-tailed)	.	.001	.064
		N	669	169	169
	Komponen Sulawesi Selatan 1	Correlation Coefficient	.258	1.000	.000
		Sig. (2-tailed)	.001	.	.996
		N	169	169	169
	Komponen Sulawesi Selatan 2	Correlation Coefficient	.143	.000	1.000
		Sig. (2-tailed)	.064	.996	.
		N	169	169	169

Berdasarkan output di atas, nilai $\hat{\alpha}$ antara F_1 dengan total income rumah tangga adalah 0.001, dan antara F_2 dengan total income rumah tangga adalah 0.064, sehingga dengan memilih tingkat signifikansi 0.150, maka dapat disimpulkan bahwa terdapat korelasi antara indikator sasaran (F_1 dan F_2) tingkat kemiskinan dengan total income rumah tangga atau total income rumah tangga untuk provinsi Sulawesi Selatan tidak termanipulasi sehingga total income rumah tangga bisa digunakan untuk mengukur tingkat kemiskinan di provinsi Sulawesi Selatan.

4.2.4 Maluku

Untuk provinsi Maluku, dengan menggunakan 167 rumah tangga, diperoleh hasil sebagai berikut:

1. Diperoleh dua komponen utama yang mencakup 67.16% total variansi dari variabel-variabel proxi.
2. Komponen pertama (F_1) mencakup 46.28% total variansi dari variabel-variabel proxi, sedangkan komponen kedua (F_2) mencakup 20.88%.
3. Kuantifikasi optimal pada setiap kategori untuk setiap variabel adalah sebagai berikut:

Tabel 4.10 Kuantifikasi optimal untuk variabel-variabel proxi provinsi Maluku

Variabel	Kategori	Kuantifikasi optimal
Jenis lantai terluas	Bambu/tanah	-0.64
	Keramik/semen/kayu jelek	-0.24
	Keramik/semen/kayu bagus	2.92
Jenis dinding terluas	Bambu/rumbia	0.00
	Tembok/kayu kondisi jelek	-0.42
	Tembok/kayu kondisi bagus	2.43
Jenis atap	Rumbia/ilalang/jerami	-1.18
	Sirap/genteng/seng/asbes kondisi jelek	-0.37
	Sirap/genteng/seng/asbes kondisi bagus	2.8
Penggunaan jamban/kakus	Bersama/umum	-0.99
	Sendiri	1.02
Sumber penerangan utama	Bukan listrik	-1.72
	Listrik tanpa meteran	-1.28
	Listrik dengan meteran	0.62
Bahan bakar utama	Kayu/arang	-0.84
	Minyak tanah	-1.21
	Gas/listrik	0.00

4. Didapat korelasi antara dua komponen dengan variabel-variabel proxi adalah sebagai berikut:

Tabel 4.11 Korelasi antara F_1 (komponen pertama) dan F_2 (komponen kedua) dengan variabel-variabel proxi untuk provinsi Maluku

Variabel	F_1	F_2
Jenis lantai terluas	0.800	0.338
Jenis dinding terluas	0.803	0.344
Jenis atap	0.822	0.270
Penggunaan jamban	0.547	-0.589
Sumber penerangan	0.413	-0.737
Bahan bakar	0.589	-0.241

Berdasarkan output di atas, komponen pertama berkorelasi cukup besar dengan variabel proxi jenis lantai terluas (0.800), jenis dinding terluas (0.803), jenis atap (0.822), dan bahan bakar utama (0.589). Sedangkan komponen kedua berkorelasi cukup besar dengan variabel proxi penggunaan jamban (-0.589), dan sumber penerangan utama (-0.737).

Oleh karena itu, tingkat kemiskinan di provinsi Maluku dapat diukur dengan indikator sasaran:

- ✓ F_1 yang lebih menekankan pada variabel jenis lantai terluas, jenis dinding terluas, jenis atap, dan bahan bakar utama

dan atau

- ✓ F_2 yang lebih menekankan pada variabel penggunaan jamban dan sumber penerangan utama

F_1 dan F_2 diperoleh dengan formula PRINCALS, yaitu sebagai berikut:

F_1 = (1/6)(0.8jenis_lantai + 0.803jenis_dinding + 0.822jenis_atap + 0.547penggunaan_jamban + 0.413penerangan + 0.589bahan_bakar)

$F_2 =$ (1/6)(0.338jenis_lantai + 0.344jenis_dinding + 0.270jenis_atap - 0.589penggunaan_jamban - 0.737penerangan - 0.241bahan_bakar)

Untuk analisis selanjutnya, pengkategorian kemiskinan provinsi Maluku dapat dilakukan berdasarkan indikator sasaran F_1 dan atau F_2.

5. Diperoleh korelasi antara indikator sasaran (F_1 dan F_2) untuk provinsi Maluku dengan total income rumah tangga

Tabel 4.12 Output uji korelasi Spearman's rho antara indikator sasaran tingkat kemiskinan dengan total income rumah tangga provinsi Maluku

Correlations

			Total income rumah tangga rata-rata dalam seminggu	Komponen Maluku 1	Komponen Maluku 2
Spearman's rho	Total income rumah tangga rata-rata dalam seminggu	Correlation Coefficient	1.000	.085	.138
		Sig. (2-tailed)	.	.275	.075
		N	669	167	167
	Komponen Maluku 1	Correlation Coefficient	.085	1.000	-.004
		Sig. (2-tailed)	.275	.	.964
		N	167	167	167
	Komponen Maluku 2	Correlation Coefficient	.138	-.004	1.000
		Sig. (2-tailed)	.075	.964	.
		N	167	167	167

Berdasarkan output di atas, nilai $\hat{\alpha}$ antara F_1 dengan total income rumah tangga adalah 0.275, dan antara F_2 dengan total income rumah tangga adalah 0.075, sehingga dengan memilih tingkat signifikansi 0.150, maka dapat disimpulkan bahwa tidak ada korelasi antara indikator sasaran (F_1) dengan total income rumah tangga, tetapi ada korelasi antara indikator sasaran (F_2) dengan total income rumah tangga.

Walaupun F_2 berkorelasi dengan total income rumah tangga, tetapi F_1 tidak, di mana F_1 menekankan pada empat variabel proxi, yaitu jenis lantai terluas, jenis dinding terluas, jenis atap, dan bahan bakar utama, lebih banyak daripada variabel-variabek proxi yang ditekankan oleh F_2, sehingga secara umum, dapat disimpulkan bahwa total income rumah tangga termanipulasi sehingga tidak bisa digunakan untuk mengukur tingkat kemiskinan di provinsi Maluku.

4.3 Kesimpulan Analisis Data

Berikut ini diberikan perbandingan hasil analisis data untuk keempat provinsi.

Tabel 4.13 Kesimpulan dari hasil analisis data

Provinsi	Variabel proxi yang ditekankan oleh F_1	Variabel proxi yang ditekankan oleh F_2	Hasil analisis untuk total income
Kalimantan Barat	Jenis atap, penggunaan jamban, sumber penerangan	Jenis lantai, jenis dinding, bahan bakar	Tidak termanipulasi (total income bisa digunakan untuk mengukur kemiskinan)
Nusa Tenggara Barat	Jenis lantai, jenis dinding, jenis atap, sumber penerangan	Penggunaan jamban, bahan bakar	Termanipulasi (total income tidak bisa digunakan untuk mengukur kemiskinan)
Sulawesi Selatan	Jenis lantai, jenis dinding, jenis atap, bahan bakar	Penggunaan jamban, sumber penerangan	Tidak termanipulasi (total income bisa digunakan untuk mengukur kemiskinan)
Maluku	Jenis lantai, jenis dinding, jenis atap, bahan bakar	Penggunaan jamban, sumber penerangan	Termanipulasi (total income tidak bisa digunakan untuk mengukur kemiskinan)

BAB 5
PENUTUP

Bab ini berisi kesimpulan yang didapat dari hasil pembahasan pada bab-bab sebelumnya serta saran untuk penelitian selanjutnya.

5.1 Kesimpulan

Kesimpulan yang dapat diperoleh dari penulisan tugas akhir ini adalah sebagai berikut:

- ✓ Metode PRINCALS dapat digunakan dalam *proxy means test* untuk membentuk indikator sasaran berdasarkan variabel-variabel proxi, di mana variabel-variabel proxi tersebut merupakan variabel kategorik.
- ✓ Indikator sasaran yang dibentuk berdasarkan metode PRINCALS dipilih satu atau lebih dari komponen-komponen utama yang telah terbentuk sesuai kebutuhan peneliti.
- ✓ Indikator sasaran yang terbentuk bisa digunakan untuk mewakili variabel sasaran untuk analisis lebih lanjut

5.2 Saran

- ✓ Untuk penulisan selanjutnya, dapat dijelaskan tentang metode PRINCALS yang memperhatikan sifat keordinalan dari variabel kategorik.
- ✓ Untuk penulisan selanjutnya, dapat dijelaskan tentang metode-metode analisis multivariat selain PRINCALS yang menggunakan teknik *optimal scaling*, seperti analisis diskriminan dengan teknik *optimal scaling*, dan korelasi kanonik dengan teknik *optimal scaling* (OVERALS).

DAFTAR PUSTAKA

Anton, Howard. 2000. Dasar-dasar Aljabar Linear. Batam: Interaksara.

De Leeuw, J and Mair, Patrick. 2009. *Gifi Methods for Optimal Scaling in R: The Package Homals*. Journal of Statistical Software, Vol. 31, Issue 4.

Dillon, William. R and Goldstain, Matthew. 1984. *Multivariate Analysis Methods and Application*. New York: John Wiley & Sons, Inc.

Gifi, Alfred. 1990. *Nonlinear Multivariate Analysis*. Chichester: Wiley.

Harville, David A. 1997. *Matrix Algebra from a Statistician's Perspective*. New York: Springer.

Horn, Roger A. & Johnson, Charles R. 1985. *Matrix Analysis*. New York: Cambridge University Press.

Jolliffe, I.T. 2004. *Principal Component Analysis Second Edition*. New York: Springer

Lecturas de Economica. 2002. *Proxy Means Test Index for Targeting Social Program: Two Methodologies and Empirical Evidence*. Lect. Econ. No. 56.

Manisera, Marica, var de Koij, Anita J, and Dussedorp, Elise. 2010. *Identifying The Component Structure of Satisfaction Scales by Nonlinear Principal Component Analysis*. Quality Technology and Quantitative Management, Vol.7, No.2, pp. 97-115.

Michailidis G and de Leeuw J. 1998. *The Gifi System of Descriptive Multivariate Analysis*. Statistical Science, 13, 307-336.

Rencher, Alvin C. 2002. *Metohds of Multivariate Analysis*. United States of America: John Wiley & Sons, Inc.

SPSS 15.0 Command Syntax Reference. 2006. PRINCALS. SPSS Inc. United States of America. pp. 1356-1364.

William D. Perrault, Jr and Forrest W. Young. 1980. *Alternating Least Square Optimal Scaling Analysis of Nonmetric Data in Marketing Research*. Journal of Marketing Research, Vol. 17, pp. 1-13

Young, Forrest W. 1981. *Quantitative of Qualitative Data*. Psychometrika, Vol. 46, No. 4 December, 357-388.

Young, Takane, and de Leeuw. 1978. *The Principal Component of Mixed Measurement Level Multivariate Data: An alternating Least Square Method with Optimal Scaling Features*. Psychometrika, Vol. 43, No. 2.

LAMPIRAN

Lampiran 1

Sifat-sifat matriks yang mendukung subbab 2.1

Definisi:

Misalkan **A** adalah matriks berukuran (n x n), $\boldsymbol{A} = \{\alpha_{ii}\}$untuk i, j = 1, 2, …, n, maka

- $tr(\boldsymbol{A}) = \sum_{i=1}^{p} \alpha_{ii}$
- **A** disebut simetri jika $\mathbf{A}^t = \mathbf{A}$, dengan $\alpha_{ij} = \alpha_{ji}$
- **A** disebut ortogonal jika $\mathbf{AA}^t = \mathbf{I}$, yaitu $\mathbf{A}^{-1} = \mathbf{A}^t$
- Untuk matriks **A**, jika terdapat bilangan skalar λ dan vektor **α**, sehingga memenuhi $\mathbf{A}\boldsymbol{\alpha} = \lambda\boldsymbol{\alpha}$ atau $(\mathbf{A} - \lambda\mathbf{I})\boldsymbol{\alpha} = 0$, maka λ disebut sebagai nilai eigen dari matriks **A** dan **α** disebut sebagai vektor eigen yang bersesuaian λ, di mana λ diperoleh dari $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$
- **A** dikatakan dapat didiagonalkan jika terdapat suatu matriks **P** sedemikian sehingga $\mathbf{P}^{-1}\mathbf{AP}$ adalah suatu matriks diagonal, matriks **P** dikatakan mendiagonalkan **A**.
- Vektor $\boldsymbol{\alpha}_\mathbf{i}$ dan $\boldsymbol{\alpha}_\mathbf{j}$ dikatakan vektor-vektor yang saling ortonormal jika $\boldsymbol{\alpha}_\mathbf{i}^\mathbf{t}\boldsymbol{\alpha}_\mathbf{i} = 1$, dan $\boldsymbol{\alpha}_\mathbf{i}^\mathbf{t}\boldsymbol{\alpha}_\mathbf{j} = 1$

Sifat-sifat matriks:

- Sifat 1 (Teorema Schur)
  Misalkan **A** adalah matriks berukuran (n x n), maka terdapat matriks ortogonal **U** berukuran (n x n), sedemikian sehingga
  $$\boldsymbol{U}^t\boldsymbol{A}\boldsymbol{U} = \boldsymbol{V}$$
  dimana **V** merupakan matriks segitiga (segitiga atas ataupun segitiga bawah), dengan elemen diagonalnya merupakan nilai eigen dari **A**. (Horn and Johnson, 1985)

- Sifat 2

  Misalkan, **A** adalah matriks yang simetris berukuran (n x n), maka terdapat suatu matriks orthogonal **U** berukuran (n x n), sedemikian sehingga $\mathbf{U}^t\mathbf{AU}$

= **D**, dimana **D** merupakan matriks diagonal yang elemen diagonalnya merupakan nilai eigen dari **A** atau dengan kata lain **A** dapat didiagonalkan secara orthogonal.

Bukti :

Karena **A** merupakan matriks berukuran (n x n), berdasarkan teorema Schur terdapat matriks ortogonal **U**, sedemikian sehingga

$$\boldsymbol{U^tAU = D}$$

dimana **U** merupakan matriks segitiga (segitiga atas ataupun segitiga bawah), dengan elemen diagonalnya merupakan nilai eigen dari **A**.

Selanjutnya, akan dibuktikan bahwa **D** merupakan matriks diagonal.

$$\begin{aligned}\boldsymbol{D^t} &= (\boldsymbol{U^tAU})^{\boldsymbol{t}}\\ &= \boldsymbol{U^tA^tU}\end{aligned}$$

Karena **A** merupakan matriks yang simetri, maka $\mathbf{A}^t = \mathbf{A}$

$$\begin{aligned}\boldsymbol{D^t} &= \boldsymbol{U^tA^tU}\\ &= \boldsymbol{U^tAU}\\ &= \boldsymbol{D}\end{aligned}$$

Karena $\mathbf{D}^t = \mathbf{D}$, dan **D** merupakan matriks segitiga, maka **D** matriks diagonal yang elemen diagonalnya merupakan nilai eigen dari matriks **A**. Karena **U** merupakan matriks ortogonal, maka $\mathbf{U}^{-1} = \mathbf{U}^t$, sehingga **U** dikatakan dapat mendiagonalkan **A** secara ortogonal.

- Sifat 3

Jika **A** merupakan matriks simetri berukuran (n x n), maka jumlah vektor eigen berbeda dari **A** adalah sebanyak n, dan vektor-vektor eigen dari **A** merupakan vektor ortonormal.

Bukti :

Berdasarkan sifat 2, karena **A** merupakan matriks simetri, maka terdapat matriks ortogonal **U** yang mendiagonalkan **A**, sedemikian sehingga $\boldsymbol{U^tAU = D}$, dimana **D** adalah matriks diagonal yang elemen diagonalnya adalah nilai eigen dari **A**. Karena $\boldsymbol{U^tAU = D}$, dan $\boldsymbol{UU^t = I}$, maka dapat diperoleh $\boldsymbol{AU = UD}$.

Sebut, $\boldsymbol{U} = (\boldsymbol{u_1} \quad \boldsymbol{u_2} \quad \cdots \quad \boldsymbol{u_n})$, di mana $\mathbf{u}_i$ merupakan vektor kolom dengan n komponen. Misalkan, λ_i adalah nilai eigen dari **A** untuk i = 1, 2, …, p, maka **D** dapat ditulis sebagai berikut.

$$\boldsymbol{D} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

Karena $\boldsymbol{AU} = \boldsymbol{UD}$, maka

$$\boldsymbol{A}(\boldsymbol{u_1} \quad \boldsymbol{u_2} \quad \cdots \quad \boldsymbol{u_n}) = (\boldsymbol{u_1} \quad \boldsymbol{u_2} \quad \cdots \quad \boldsymbol{u_n}) \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

$$(\boldsymbol{Au_1} \quad \boldsymbol{Au_2} \quad \cdots \quad \boldsymbol{Au_n}) = (\lambda_1\boldsymbol{u_1} \quad \lambda_2\boldsymbol{u_2} \quad \cdots \quad \lambda_n\boldsymbol{u_n})$$

Vektor $\mathbf{u}_i$ dan λ_i memenuhi persamaan $\mathbf{Au}_i = \lambda_i\mathbf{u}_i$, maka $\mathbf{u}_i$ merupakan vektor eigen dari matriks **A**, yang bersesuaian dengan nilai eigen λ_i, untuk setiap i = 1, 2, …, p.

$$\boldsymbol{U}^t\boldsymbol{U} = \begin{pmatrix} \boldsymbol{u_1^t} \\ \boldsymbol{u_2^t} \\ \vdots \\ \boldsymbol{u_n^t} \end{pmatrix} (\boldsymbol{u_1} \quad \boldsymbol{u_2} \quad \cdots \quad \boldsymbol{u_n}) = \begin{pmatrix} \boldsymbol{u_1^t u_1} & \boldsymbol{u_1^t u_2} & \dots & \boldsymbol{u_1^t u_n} \\ \boldsymbol{u_2^t u_1} & \boldsymbol{u_2^t u_2} & \dots & \boldsymbol{u_2^t u_n} \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{u_n^t u_1} & \boldsymbol{u_n^t u_2} & \dots & \boldsymbol{u_n^t u_n} \end{pmatrix} = \boldsymbol{I}$$

Karena $\boldsymbol{u_i^t u_i} = 1$, dan $\boldsymbol{u_i^t u_j} = 0$, maka terbukti bahwa **A** memiliki n vektor eigen yang merupakan vektor ortonormal.

- Sifat 4

  $tr(\boldsymbol{AB}) = tr(\boldsymbol{BA})$ sehingga $\boldsymbol{tr}(\boldsymbol{ABC}) = \boldsymbol{tr}(\boldsymbol{CAB}) = \boldsymbol{tr}(\boldsymbol{BCA})$

  Bukti :

  Misalkan, **A** berukuran (n x c) dan **B** berukuran (c x n). Misalkan tr(**AB**) = tr(**P**) = $\sum_{i=0}^{n} p_{ii}$ dan tr(**BA**) = **T** = $\sum_{j=0}^{c} t_{jj}$.

$$\text{tr}(\mathbf{AB}) = \sum_{i=0}^{n} p_{ii} = \sum_{i=1}^{n}\left(\sum_{j=1}^{c} a_{ij}\, b_{ji}\right) = \sum_{i=1}^{n}\left(\sum_{j=1}^{c} b_{ji}\, a_{ij}\right)$$

$$= \sum_{j=1}^{c}\left(\sum_{i=1}^{n} b_{ji}\, a_{ij}\right) = \sum_{j=0}^{c} t_{jj} = \text{tr}(\mathbf{BA})$$

Karena tr(**AB**) = tr(**BA**), maka tr(**ABC**) = tr(**BCA**) = tr(**CAB**)

Lampiran 2

Misalkan, $X_1, X_2, \dots, X_p$ adalah suatu variabel random. Misalkan, $\boldsymbol{x_1}, \boldsymbol{x_2}, \dots, \boldsymbol{x_p}$ adalah n pengamatan untuk $X_1, X_2, \dots, X_p$. Misalkan, $\bar{x}_j$ adalah mean sampel dari variabel X_j, di mana j = 1, 2, …, p. Misalkan, s_{jj} adalah variansi sampel dari X_j, sehingga $s_{jj} = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_{ij} - \bar{x}_j\right)^2$. Misalkan, s_{jk} adalah kovariansi antara X_j dan X_k, sehingga $s_{jk} = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_{ij} - \bar{x}_j\right)(x_{ik} - \bar{x}_k)$. Misalkan, r_{jk} adalah korelasi antara X_j dan X_k, sehingga $r_{jk} = \frac{s_{jk}}{\sqrt{s_{jj}}\sqrt{s_{kk}}}$.

Sebut, $\boldsymbol{x_j} *$ sebagai vektor yang elemen-elemennya adalah elemen-elemen $\boldsymbol{x_j}$ yang telah distadarisasi, maka

$$\boldsymbol{x_j} * = \begin{pmatrix} \frac{x_{1j} - \bar{x}_j}{\sqrt{s_{jj}}} \\ \frac{x_{1j} - \bar{x}_j}{\sqrt{s_{jj}}} \\ \vdots \\ \frac{x_{nj} - \bar{x}_j}{\sqrt{s_{jj}}} \end{pmatrix}$$

Sebut, $\boldsymbol{X} * = (\boldsymbol{x_1} * \quad \boldsymbol{x_2} * \quad \cdots \quad \boldsymbol{x_p} *)$.

Misalkan, $\boldsymbol{R}$ adalah matriks korelasi antara $X_1, X_2, \dots, X_p$. Akan dibuktikan jika $\boldsymbol{R} = \frac{\boldsymbol{1}}{\boldsymbol{n-1}}\boldsymbol{X^t} * \boldsymbol{X} *$.

Bukti :

$$\boldsymbol{X^t} * \boldsymbol{X} * = \begin{pmatrix} \boldsymbol{x_1^t} * \\ \boldsymbol{x_2^t} * \\ \vdots \\ \boldsymbol{x_p^t} * \end{pmatrix} (\boldsymbol{x_1} * \quad \boldsymbol{x_2} * \quad \cdots \quad \boldsymbol{x_p} *)$$

$$= \begin{pmatrix} \boldsymbol{x_1^t} * \boldsymbol{x_1} * & \boldsymbol{x_1^t} * \boldsymbol{x_2} * & \cdots & \boldsymbol{x_1^t} * \boldsymbol{x_p} * \\ \boldsymbol{x_2^t} * \boldsymbol{x_1} * & \boldsymbol{x_2^t} * \boldsymbol{x_2} * & \cdots & \boldsymbol{x_2^t} * \boldsymbol{x_p} * \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{x_p^t} * \boldsymbol{x_1} * & \boldsymbol{x_p^t} * \boldsymbol{x_2} * & \cdots & \boldsymbol{x_p^t} * \boldsymbol{x_p} * \end{pmatrix}$$

$$= \begin{pmatrix} \frac{\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)^2}{s_{11}} & \frac{\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)(x_{i2}-\bar{x}_2).}{\sqrt{s_{11}}\sqrt{s_{22}}}. & \cdots & \frac{\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)(x_{ip}-\bar{x}_p).}{\sqrt{s_{11}}\sqrt{s_{pp}}}. \\ \frac{\sum_{i=1}^{n}(x_{i2}-\bar{x}_2)(x_{i1}-\bar{x}_1).}{\sqrt{s_{22}}\sqrt{s_{11}}}. & \frac{\sum_{i=1}^{n}(x_{i2}-\bar{x}_2)^2}{s_{22}} & \cdots & \frac{\sum_{i=1}^{n}(x_{i2}-\bar{x}_2)(x_{ip}-\bar{x}_p).}{\sqrt{s_{22}}\sqrt{s_{pp}}}. \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\sum_{i=1}^{n}(x_{ip}-\bar{x}_p)(x_{i1}-\bar{x}_1).}{\sqrt{s_{pp}}\sqrt{s_{11}}}. & \frac{\sum_{i=1}^{n}(x_{ip}-\bar{x}_p)(x_{i2}-\bar{x}_2).}{\sqrt{s_{pp}}\sqrt{s_{22}}}. & \dots & \frac{\sum_{i=1}^{n}(x_{ip}-\bar{x}_p)^2}{s_{pp}} \end{pmatrix}$$

$$\frac{\mathbf{1}}{\boldsymbol{n-1}}\boldsymbol{X^t} * \boldsymbol{X} * =$$

$$= \frac{\mathbf{1}}{\boldsymbol{n-1}} \begin{pmatrix} \frac{\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)^2}{s_{11}} & \frac{\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)(x_{i2}-\bar{x}_2).}{\sqrt{s_{11}}\sqrt{s_{22}}}. & \cdots & \frac{\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)(x_{ip}-\bar{x}_p).}{\sqrt{s_{11}}\sqrt{s_{pp}}}. \\ \frac{\sum_{i=1}^{n}(x_{i2}-\bar{x}_2)(x_{i1}-\bar{x}_1).}{\sqrt{s_{22}}\sqrt{s_{11}}}. & \frac{\sum_{i=1}^{n}(x_{i2}-\bar{x}_2)^2}{s_{22}} & \cdots & \frac{\sum_{i=1}^{n}(x_{i2}-\bar{x}_2)(x_{ip}-\bar{x}_p).}{\sqrt{s_{22}}\sqrt{s_{pp}}}. \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\sum_{i=1}^{n}(x_{ip}-\bar{x}_p)(x_{i1}-\bar{x}_1).}{\sqrt{s_{pp}}\sqrt{s_{11}}}. & \frac{\sum_{i=1}^{n}(x_{ip}-\bar{x}_p)(x_{i2}-\bar{x}_2).}{\sqrt{s_{pp}}\sqrt{s_{22}}}. & \dots & \frac{\sum_{i=1}^{n}(x_{ip}-\bar{x}_p)^2}{s_{pp}} \end{pmatrix}$$

$$= \begin{pmatrix} \frac{s_{11}}{s_{11}} & \frac{s_{12}}{\sqrt{s_{11}}\sqrt{s_{22}}}. & \cdots & \frac{s_{1p}}{\sqrt{s_{11}}\sqrt{s_{pp}}} \\ \frac{s_{21}}{\sqrt{s_{22}}\sqrt{s_{11}}}. & \frac{s_{22}}{s_{22}} & \cdots & \frac{s_{2p}}{\sqrt{s_{22}}\sqrt{s_{pp}}}. \\ \vdots & \vdots & \ddots & \vdots \\ \frac{s_{p1}}{\sqrt{s_{pp}}\sqrt{s_{11}}}. & \frac{s_{p2}}{\sqrt{s_{pp}}\sqrt{s_{22}}}. & \dots & \frac{s_{pp}}{s_{pp}} \end{pmatrix}$$

$$= \begin{pmatrix} 1 & r_{12} & \cdots & r_{1p} \\ r_{21} & 1 & \cdots & r_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{p1} & r_{12} & \dots & 1 \end{pmatrix} = \boldsymbol{R}$$

Terbukti bahwa $\boldsymbol{R} = \frac{\mathbf{1}}{\boldsymbol{n-1}}\boldsymbol{X^t} * \boldsymbol{X} *$

Lampiran 3

Metode *Gram-Schmidt*

Misalkan, $\boldsymbol{v_1}, \boldsymbol{v_2}, \dots, \boldsymbol{v_n}$ adalah suatu vektor sembarang dengan ukuran sama. Misalkan, $\langle \boldsymbol{v_i}, \boldsymbol{v_k} \rangle$ adalah hasil kali dalam vektor $\boldsymbol{v_i}$ dengan vektor $\boldsymbol{v_k}$ dan misalkan $\|\boldsymbol{v_i}\|$ menyatakan panjang vektor $\boldsymbol{v_i}$. Akan dicari vektor-vektor $\boldsymbol{u_1}, \boldsymbol{u_2}, \dots, \boldsymbol{u_n}$ yang saling ortogonal berdasarkan $\boldsymbol{v_1}, \boldsymbol{v_2}, \dots, \boldsymbol{v_n}$ dengan metode *Gram-Schmidt*. Langkah-langkahnya adalah sebagai berikut:

1. Anggap $\boldsymbol{u_1} = \boldsymbol{v_1}$
2. Vektor $\boldsymbol{u_2}$ dicari dengan $\boldsymbol{u_2} = \boldsymbol{v_2} - \frac{\langle \boldsymbol{u_1}, \boldsymbol{v_2} \rangle}{\|\boldsymbol{u_1}\|^2}$
3. Vektor $\boldsymbol{u_3}$ dicari dengan $\boldsymbol{u_3} = \boldsymbol{v_3} - \frac{\langle \boldsymbol{u_1}, \boldsymbol{v_3} \rangle}{\|\boldsymbol{u_1}\|^2} - \frac{\langle \boldsymbol{u_2}, \boldsymbol{v_3} \rangle}{\|\boldsymbol{u_2}\|^2}$

$$\vdots$$

4. Vektor $\boldsymbol{u_n}$ dicari dengan $\boldsymbol{u_n} = \boldsymbol{v_n} - \frac{\langle \boldsymbol{u_1}, \boldsymbol{v_n} \rangle}{\|\boldsymbol{u_1}\|^2} - \frac{\langle \boldsymbol{u_2}, \boldsymbol{v_n} \rangle}{\|\boldsymbol{u_2}\|^2} - \cdots - \frac{\langle \boldsymbol{u_{n-1}}, \boldsymbol{v_n} \rangle}{\|\boldsymbol{u_{n-1}}\|^2}$

Vektor $\boldsymbol{u_1}, \boldsymbol{u_2}, \dots, \boldsymbol{u_n}$ yang diperoleh dengan langkah di atas adalah saling ortogonal.