ANALISIS KOMPONEN UTAMA MELALUI PENDEKATAN RUANG DUAL

(Skripsi)

Oleh

SELVI ANGGRAINI

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2016

ABSTRAK

ANALISIS KOMPONEN UTAMA MELALUI PENDEKATAN RUANG DUAL

Oleh

SELVI ANGGRAINI

Analisis Komponen Utama digunakan untuk mereduksi dimensi data. Data yang berdimensi p akan direduksi menjadi berdimensi q, dengan q kurang dari p. Sehingga interpretasi dapat lebih mudah dilakukan dalam ruang berdimensi q. Dimensi data multivariate disajikan dalam ruang variabel dan ruang individu. Ruang variabel dan ruang individu dihubungkan dengan diagram dual. Ruang variabel mempunyai ruang dual, begitu pula untuk ruang individu. Melalui pendekatan ruang dual, analisis komponen utama berusaha mereduksi dimensi ruang. Pereduksian ruang variabel diawali dengan mencari bentuk bilinear dari matriks data terpusat, lalu mencari akar karakteristik dan vektor karakteristiknya hingga mendapatkan kombinasi linear yang dnamakan komponen utama. Kemudian mencari kualitas global komponen utama dari proporsi akar karakteristik dengan trace-nya.

Kata kunci : ***Multivariate, Komponen Utama, Bentuk Bilinear, Ruang Dual, Kualitas Global.***

ANALISIS KOMPONEN UTAMA MELALUI PENDEKATAN RUANG DUAL

Oleh

Selvi Anggraini

Skripsi

Sebagai Salah Satu Syarat untuk Memperoleh Gelar
SARJANA SAINS

Pada

Jurusan Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2016

Judul Skripsi : **ANALISIS KOMPONEN UTAMA MELALUI PENDEKATAN RUANG DUAL**

Nama Mahasiswa : Selvi Anggraini

Nomor Pokok Mahasiswa : 1217031061

Jurusan : Matematika

Fakultas : Matematika dan Ilmu Pengetahuan Alam

MENYETUJUI

1. Komisi Pembimbing

Rudi Ruswandi, M.Si.
NIP 19560208 198902 1 001

Amanto, M.Si.
NIP 19730314 200012 1 002

2. Ketua Jurusan Matematika

Drs. Tiryono Ruby, M.Sc., Ph.D.
NIP 19620704 198803 1 002

MENGESAHKAN

1. Tim Penguji

Ketua : **Rudi Ruswandi, M.Si.** ....

Sekretaris : **Amanto, M.Si.** ....

Penguji
Bukan Pembimbing : **Widiarti, M.Si.** ....

2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam

Prof. Warsito, S.Si., D.E.A., Ph.D.
NIP 19710212 199512 1 001

Tanggal Lulus Ujian Skripsi : **10 Juni 2016**

PERNYATAAN

Saya yang bertandatangan di bawah ini, menyatakan bahwa skripsi saya yang berjudul **"ANALISIS KOMPONEN UTAMA MELALUI PENDEKATAN RUANG DUAL "** merupakan hasil karya saya sendiri dan bukan hasil karya orang lain. Semua hasil tulisan yang tertuang dalam skripsi ini telah mengikuti kaidah penulisan karya ilmiah Universitas Lampung. Apabila dikemudian hari terbukti bahwa skripsi ini merupakan hasil salinan atau dibuat oleh orang lain, maka saya bersedia menerima sanksi sesuai dengan ketentuan akademik yang berlaku.

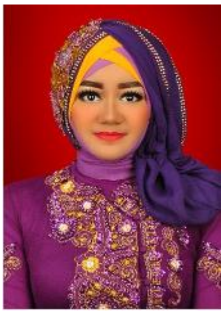
Bandar Lampung, Juni 2016

Penulis

Selvi Anggraini

NPM. 1217031061

RIWAYAT HIDUP



Penulis dilahirkan di Padang Bindu, Baturaja Sumatera Selatan pada tanggal 23 Februari 1995. Penulis merupakan anak pertama dari dua bersaudara, dari pasangan bapak Dariza dan ibu Nur'aimah, serta kakak dari Bayu Andika.

Penulis memulai pendidikan Sekolah Dasar di SDN 02 Bumi Pratama Mandira Sumatera Selatan pada tahun 2000-2006. Sekolah Menengah Pertama di SLTP Budi Pratama Sumatera Selatan pada tahun 2006-2009. Kemudian Sekolah Menengah Atas di SMA Bina Dharma Mandira, Ogan Komering Ilir Sumatera Selatan pada tahun 2009-2012.

Tahun 2012, penulis terdaftar sebagai mahasiswa Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung melalui jalur Seleksi Nasional Masuk Perguruan Tinggi Negeri (SNMPTN) tertulis. Selama menjadi mahasiswa penulis pernah menjadi asisten praktikum Pengantar Teknologi Informasi dan asisten praktikum Algoritma dan Pemrograman di Jurusan Matematika FMIPA Unila, asisten Statistika di Jurusan Biologi FMIPA Unila, asisten Matematika (Kalkulus) di Program Studi D3 MI dan S1 Ilmu Komputer Jurusan Ilmu Komputer FMIPA Unila. Penulis juga aktif dalam kegiatan kemahasiswaan tingkat jurusan, fakultas, universitas, maupun nasional.

Penulis tercatat pernah menjabat sebagai Bendahara Gematika periode 2012-2013, AMAR periode 2012-2013, Garuda BEM FMIPA periode 2012-2013, Bendahara PANSUS 2013, Wakil Bendahara Umum HIMATIKA periode 2013-2014, Bendahara Umum HIMATIKA periode 2014-2015, Dewan Pembina Organisasi HIMATIKA sejak September 2015. Pada bulan Februari tahun 2015, penulis melakukan Kerja Praktik (KP) di Badan Pusat Statistik (BPS) Kota Bandar Lampung. Pada bulan Agustus tahun 2015, penulis melakukan Kuliah Kerja Nyata (KKN) serta dipercaya mewakili Universitas Lampung untuk bergabung dalam KKN Kebangsaan 2015 sebagai Sekretaris Koordinator Desa yang diabdikan di Desa Sungai Kayu Ara, Kecamatan Sungai Apit, Kabupaten Siak, Provinsi Riau.

Selama menjadi mahasiswa di Universitas Lampung, penulis juga selalu mendapatkan Beasiswa setiap tahunnya hingga lulus. Penulis tercatat pernah mendapatkan Beasiswa Bantuan Belajar Mahasiswa (BBM) pada tahun akademik 2013/2014, lalu Beasiswa Peningkatan Prestasi Akademik (PPA) pada tahun akademik 2014/2015, dan Beasiswa Bantuan Biaya Pendidikan Peningkatan Prestasi Akademik (BBP-PPA) pada tahun akademik 2015/2016.

M O T T O

Wahai orang-orang yang beriman, bersabarlah kamu dan kuatkanlah kesabaranmu dan tetaplah bersiap-siaga dan bertaqwalah kepada Allah, supaya kamu menang

(Q.S. Ali Imran : 200)

"Kami menyampaikan kabar gembira kepadamu dengan benar. Maka janganlah kamu termasuk orang-orang yang berputus asa"

(Q.S Al-Hijr : 55)

Manisnya keberhasilan akan menghapus pahitnya kesabaran. Nikmatnya kemenangan melenyapkan letihnya perjuangan. Menuntaskan pekerjaan dengan baik akan melenyapkan lelahnya jerih payah

(Dr. Aidh Al Zarni)

Jangan menunggu bisa baru mau melakukannya, tapi segera lakukanlah ! Kamu pasti bisa !

(Selvi Anggraini)

P E R S E M B A H A N

Dengan penuh rasa syukur kepada Allah SWT dan dengan kerendahan hati saya persembahkan karya kecil dan sederhana ini sebagai tanda bakti dan cinta kepada semua orang yang senantiasa mendukung dan dengan tulus mendoakan kelancaran demi terciptanya karya ini.

Abah, Umak, serta Adek Bayu yang sangat saya sayangi, yang dengan tulus selalu memberikan semangat serta dukungan dan doa yang amat sangat luar biasa demi keberhasilan penulis.

Sahabat-sahabat terbaik yang selalu ada. Terimakasih atas kebersamaan, keceriaan, semangat serta motivasi yang diberikan kepada penulis.

Almamaterku tercinta --- Universitas Lampung

SANWACANA

Alhamdulillahirobbil'alamin, segala puji dan syukur kepada Allah SWT atas izin serta ridho-Nya penulis dapat menyelesaikan skripsi yang berjudul **"Analisis Komponen Utama melalui Pendekatan Rual Dual"**. Shalawat teriring salam kepada junjungan nabi besar Muhammad SAW yang telah menjadi suri tauladan yang baik bagi kita semua pengikutnya.

Pada proses penyusunan skripsi ini, penulis memperoleh banyak bimbingan, kritik, dan saran yang membangun sehingga skripsi ini mampu penulis selesaikan. Oleh karena itu, penulis ingin mengucapkan terimakasih kepada :

1. Bapak Rudi Ruswandi, M.Si., selaku pembimbing pertama yang senantiasa membimbing dan memberikan arahan kepada penulis dalam menyelesaikan skripsi ini.

2. Bapak Amanto, M.Si., selaku pembimbing kedua yang telah banyak membantu dan selalu sabar memberikan pengarahan, serta dukungan dalam proses penyusunan skripsi ini.

3. Ibu Widiarti, M.Si., selaku penguji yang telah memberikan kritik, saran, dan masukan yang membangun kepada penulis.

4. Bapak Dr. Muslim Ansori, M.Si., selaku pembimbing akademik yang tak pernah letih memberikan masukan dan nasihat selama penulis menjalani proses perkuliahan.

5. Bapak Drs. Tiryono Ruby, M.Sc., Ph.D., selaku Ketua Jurusan Matematika FMIPA Universitas Lampung.
6. Bapak Prof. Warsito, S.Si.,D.E.A.,Ph.D., selaku Dekan FMIPA Universitas Lampung.
7. Seluruh dosen, staff, dan karyawan Jurusan Matematika FMIPA Universitas Lampung yang telah memberikan ilmu pengetahuan dan bantuan kepada penulis.
8. Untuk kedua orang tua yang menjadi semangat hidupku “Abah dan Umak”, serta adik kecilku “Bayu”, terimakasih telah menjadi semangat yang tak pernah hilang, pendukung yang selalu ada baik raga maupun doa.
9. Sahabat seperjuangan penulis dalam merangkai asa dan cita juga harapan Rendi Andika, Yama, Imah, Yanti, Ernia, Desti, Icha, Maya, Citra, Erni, Dwi, Anggi, Anwar, Taufik, Rendi.R, Gery, Pras, Candra.
10. Keluarga Besar, Seluruh Pimpinan dan Anggota serta Gematika HIMATIKA FMIPA Universitas Lampung periode 2014-2015.
11. Teman-teman Matematika angkatan 2012 atas kebersamaannya selama ini.
12. Kawan-kawan #KBN48 Kuliah Kerja Nyata (KKN) Kebangsaan 2015.
13. Seluruh pihak yang telah membantu dalam penyusunan skripsi ini yang tidak dapat disebutkan satu persatu.

Bandar Lampung, Juni 2016
Penulis,

Selvi Anggraini

DAFTAR ISI

Halaman

DAFTAR TABEL

I. PENDAHULUAN

1.1 Latar Belakang dan Masalah

Dalam statistika dikenal dua jenis data yaitu data *univariate* dan data *multivariate*. Data *univariate* diperoleh cukup dengan memperhatikan satu peubah atau karakter saja dari satu individu. Sedangkan data *multivariate* diperoleh dengan memperhatikan dua bahkan lebih banyak lagi peubah yang merupakan karakter dari individu yang sama. Secara umum, karena memperhatikan banyak karakter dari individu yang sama, maka akan ada saling ketergantungan atau keterkaitan antar karakter yang diamati tersebut atau dikenal juga dengan istilah saling berkorelasi. Keadaan tersebut yang dapat menjadi fitur yang dapat membedakan antara peubah ganda (*multivariate*) dan peubah tunggal (*univariate*).

Salah satu teknik statistika yang banyak digunakan dalam analisis peubah ganda (*multivariate*) adalah Analisis Komponen Utama (AKU), karena analisis ini mampu mereduksi dimensi dari segugus data peubah ganda. Analisis Komponen Utama (AKU) biasanya digunakan untuk mentransformasi struktur data, yaitu untuk mengurangi dimensi himpunan peubah yang biasanya terdiri atas beberapa peubah yang saling berkorelasi menjadi peubah-peubah baru (komponen-komponen utama) yang saling tidak berkorelasi dengan dimensi yang lebih kecil dari dimensi peubah-peubah semula tanpa harus banyak kehilangan informasi (keragaman dalam himpunan data sebanyak mungkin dipertahankan).

Komponen utama dapat memberikan informasi yang besar seperti yang diterangkan oleh variabel-variabel awal. Sehingga sejumlah komponen utama tersebut dapat menggantikan sebanyak p-variabel, artinya misalkan data awal yang terdiri dari n-individu dengan p-variabel kemudian direduksi menjadi n-individu dengan beberapa komponen utama. Komponen utama tersebut terbentuk dari kombinasi linear dari variabel-variabel awal, dan di antara kombinasi linear yang terbentuk tidak akan terjadi korelasi antara yang satu dengan yang lainnya.

Dalam sebuah kasus *multivariate* yang mencakup banyak variabel dan tidak menutup kemungkinan bahwa individu yang akan diteliti juga terdiri dari jumlah yang besar, tentu menjadi tidak mudah untuk menginterpretasikan sebuah pengamatan *multivariate* terhadap individu dalam jumlah yang besar tersebut. Banyak teknik *multivariate* yang dapat digunakan dalam rangka mempermudah penginterpretasian, salah satu yang dapat digunakan adalah teknik analisis komponen utama. Untuk mendapatkan komponen utama yang diinginkan tersebut maka akan dilakukan reduksi dimensi dari segugus data berdimensi besar menjadi data yang berdimensi kecil. Dalam hal ini penulis tertarik untuk melakukan penelitian terkait analisis komponen utama tersebut namun akan dilakukan melalui pendekatan ruang dual.

Ruang dual merupakan himpunan semua bentuk linear yang didefinisikan pada ruang vektor real. Sehingga dalam hal ini akan dipandang data *multivariate* yang ada sebagai vektor-vektor acak yang akan disajikan dalam bentuk matriks. Suatu bentuk linear yang merupakan anggota dari ruang dual tersebut adalah suatu pemetaan linear dengan domain ruang vektor dan range-nya berupa lapangan. Sehingga reduksi dimensi matriks data *multivariate* akan dilakukan dengan memperhatikan konsep-konsep pada ruang dual tersebut.

1.2 Perumusan Masalah

Dari latar belakang di atas, maka permasalahan yang akan dibahas dalam skripsi ini adalah mereduksi dimensi matriks yang akan dilakukan melalui pendekatan ruang dual. Dalam hal ini, bentuk linear yang merupakan anggota dari ruang dual tersebut akan memetakan dari ruang vektor real ke himpunan bilangan real.

1.3 Tujuan Penelitian

Tujuan dari penelitian ini adalah :

1. Mengkaji analisis komponen utama dengan menggunakan pendekatan ruang dual dari ruang vektor real.
2. Mendapatkan dimensi data yang lebih kecil dari sejumlah besar himpunan variabel dengan mempertahankan sebanyak mungkin keragaman dalam himpunan data tersebut.

II. TINJAUAN PUSTAKA

2.1 Konsep Dasar Matriks

2.1.1 Definisi Matriks

Matriks adalah suatu susunan bilangan berbentuk segiempat. Bilangan-bilangan dalam susunan itu disebut anggota dalam matriks tersebut (Anton, 2000).

2.1.2 Transpose Matriks

Jika **A** adalah sembarang matriks berukuran m x n, maka transpos **A** dinyatakan dengan $\mathbf{A}^{\mathbf{T}}$, didefinisikan sebagai matriks berukuran n x m yang didapatkan dengan mempertukarkan baris dan kolom dari **A** yaitu kolom pertama dari $\mathbf{A}^{\mathbf{T}}$ adalah baris pertama dari **A,** kolom kedua dari $\mathbf{A}^{\mathbf{T}}$ adalah baris kedua dari **A,** dan seterusnya (Anton, 2000).

2.1.3 Matriks Simetrik dan *Trace* Matriks

Misalkan **A** adalah sembarang matriks berukuran m x n, jika $\mathbf{A} = \mathbf{A}^{\mathbf{T}}$ maka **A** simetrik (Anton, 2000).

Jika diberikan sebarang matriks **B** berukuran n x n, maka *trace* dari matriks **B** tersebut didefinisikan sebagai jumlah dari unsur-unsur diagonalnya, dan dinotasikan dengan Tr(**B**) sehingga :

Tr(**B**) = $b_{11} + b_{22} + \ldots + b_{nn}$ (Mattjik dan Sumertajaya, 2011).

2.1.4 Determinan Matriks

Misalkan **A** adalah suatu matriks bujur sangkar. Determinan dinyatakan dengan **det,** dan selanjutnya mendefinisikan **det (A)** sebagai jumlah semua hasil kali dasar bertanda dari **A**. Nilai **det (A)** disebut **determinan A** (Anton, 2000).

2.1.5 Invers Matriks

Invers sebuah matriks bujur sangkar **A**, adalah $\mathbf{A}^{-1}$ dan didefinisikan sebagai matriks yang jika dikalikan dengan matriks asal **A** menghasilkan matriks identitas. Matriks invers selalu sebuah matriks bujur sangkar berordo sama dengan matriks asal (Anton, 2000).

Teorema 2.1.5.1.

Jika suatu matriks mempunyai invers, maka inversnya tunggal.

Bukti :

Misalkan invers dari suatu matriks **A** adalah **B** dan **C** maka :

$\mathbf{A\,B = B\,A = I}$

$\mathbf{A\,C = C\,A = I}$

Akan ditunjukkan bahwa $\mathbf{B = C}$

$\mathbf{B = B\,I = B\,A\,C = (B\,A)\,C = I\,C = C}$

Karena $\mathbf{B = C}$ maka terbukti bahwa invers dari **A** bersifat tunggal.

Teorema 2.1.5.2.

Jika **A** dan **B** adalah sebuah matriks dan **AB** mempunyai invers, maka inversnya adalah $\mathbf{(AB)^{-1} = B^{-1}\,A^{-1}}$

Bukti.

Jika $\mathbf{(AB)\,(AB)^{-1} = I}$

dan **A B B^{-1} A^{-1} = A (BB^{-1}) A^{-1} = A I A^{-1} = AA^{-1} = I**

maka benar bahwa **B^{-1} A^{-1}** adalah invers dari **AB**

2.2 Ruang Vektor

Ruang vektor adalah ruang atau himpunan dimana elemen-elemennya berupa vektor. Akan tetapi, himpunan-himpunan tersebut harus memiliki struktur yang sama, mereka dilengkapi dengan operasi penjumlahan (+) dan homotesi (.) yang sesuai dengan aturan-aturannya sebagai berikut.

(i). Sifat operasi penjumlahan

Misal P adalah ruang vektor. Operasi penjumlahan (+) adalah pemetaan dari P + P → P yaitu pemetaan dari ruang vektor yang didefinisikan suatu operasi penjumlahan dengan ruang vektor pula dan dipetakan ke ruang vektor juga, yang memetakan pasangan $(\bar{x}, \bar{y})$ menjadi $\bar{Z} = \bar{x} + \bar{y}$. Untuk setiap $\bar{x}$ dan $\bar{y}$ di P dan bersifat:

a. $\bar{x} + \bar{y} = \bar{y} + \bar{x}$ (komutatif)

b. $(\bar{x} + \bar{y}) + \bar{z} = \bar{x} + (\bar{y} + \bar{z})$ (asosiatif)

c. $\bar{x} + \bar{0} = \bar{0} + \bar{x} = \bar{x}$ (untuk setiap $\bar{x}$ di P)

d. Setiap $\bar{x}$ di P memiliki invers $\bar{y}$ sehingga $\bar{x} + \bar{y} = \bar{0}$ ($\bar{y}$ biasa ditulis -$\bar{x}$)

(ii). Sifat operasi homotesi atau perkalian dengan skalar

Misal P adalah ruang vektor dan R adalah ruang bilangan real. Homotesi (.) adalah pemetaan dari R x P → P yaitu pemetaan dari ruang bilangan real yang didefinisikan suatu operasi perkalian dengan ruang vektor dan dipetakan ke ruang vektor pula, yang memetakan setiap pasangan (α, $\bar{x}$) ; α real dan $\bar{x}$ di P, menjadi

$\alpha\ .\ \bar{x}$ (atau $\alpha\bar{x}$). Dengan R menyatakan himpunan bilangan real, homotesi (.) memiliki sifat :

a. $1\ .\ \bar{x} = \bar{x}$; 1 di R
b. $\alpha\ (\beta\ \bar{x}) = (\alpha\ \beta)\ \bar{x}$; α , β di R
c. $0\ .\ \bar{x} = \bar{0}$; 0 di R dan $\bar{0}$ di P
d. $\alpha\ (\bar{x} + \bar{y}) = \alpha\ \bar{x} + \alpha\ \bar{y}$
e. $(\alpha + \beta)\ \bar{x} = \alpha\ \bar{x} + \beta\ \bar{x}$

Struktur itulah yang disebut struktur ruang vektor dan himpunan yang dilengkapi dengan struktur itu dinamakan ruang vektor. Karena homotesi pada struktur tersebut menggunakan skalar real, maka ruang vektor demikian disebut ruang vektor atas R atau ruang vektor real (Djauhari, 1988).

2.3 Kombinasi Linear

Definisi Kombinasi Linear :

Misalkan E ruang vektor real sembarang, dan B = { $\bar{b}_1, \bar{b}_2$, …, $\bar{b}_p\}$ himpunan bagian dari E. Vektor $\bar{x}$ di E adalah kombinasi linear dari B (atau dari b_1 , b_2 , …, b_p), bila terdapat bilangan-bilangan real x_1 , x_2 , …, x_p sehingga :

$$\bar{x} = x_1\ \bar{b}_1 + x_2\ \bar{b}_2 + \ldots + x_p\ \bar{b}_p$$

2.4 Bebas Linear

Definisi Bebas Linear :

Vektor-vektor $\bar{x}_1$, $\bar{x}_2$, …, $\bar{x}_k$ dikatakan bebas linear bila persamaan :

$\lambda_1\ \bar{x}_1 + \lambda_2\ \bar{x}_2 + \ldots + \lambda_k\ \bar{x}_k = \bar{0}$; dengan $\boldsymbol{\lambda}$ skalar

hanya dipenuhi oleh $\lambda_1 = \lambda_2 = \ldots = \lambda_k$ = 0 (Djauhari, 1988).

2.5 Basis

Definisi Basis :

Misalkan E ruang vektor real sembarang, dan B = { $\bar{b}_1, \bar{b}_2, \ldots, \bar{b}_p$} himpunan bagian dari E. Himpunan B dinamakan basis dari E, bila setiap vektor di E dapat ditulis secara unik sebagai kombinasi linear dari B (Djauhari, 1988).

2.5.1 Basis Kanonik

Definisi Basis Kanonik :

Misalkan E ruang vektor real sembarang, dan B = { $\bar{b}_1, \bar{b}_2, \ldots, \bar{b}_p$} basis dari E. B dinamakan basis kanonik bila :

$$\bar{b}_1 = \begin{bmatrix} 1 \\ 0 \\ . \\ . \\ 0 \end{bmatrix}$$

$$\bar{b}_2 = \begin{bmatrix} 0 \\ 1 \\ . \\ . \\ 0 \end{bmatrix}$$

.

.

.

$$\bar{b}_p = \begin{bmatrix} 0 \\ 0 \\ . \\ . \\ 1 \end{bmatrix}$$

2.6 Pemetaan Linear

Definisi Pemetaan Linear :

Pemetaan f dari ruang vektor E ke dalam ruang vektor F dinamakan pemetaan linear bila :

(i). $f(\bar{x} + \bar{y}) = f(\bar{x}) + f(\bar{y})$

(ii). $f(\lambda \bar{x}) = \lambda f(\bar{x})$

untuk setiap $\bar{x}$ dan $\bar{y}$ di E dan setiap λ di R (Djauhari, 1988).

2.7 Ruang Dual

Definisi Ruang Dual :

Misalkan L(E,R) merupakan himpunan semua pemetaan linear dari setiap vektor di E ke dalam suatu bilangan real di R sehingga L(E,R) juga merupakan ruang vektor. Secara khususnya pula, anggota dari L(E,R) dinamakan bentuk linear.

Jika f memetakan setiap vektor di E ke dalam suatu bilangan real di R maka f merupakan suatu bentuk linear jika :

$f(\bar{x} + \bar{y}) = f(\bar{x}) + f(\bar{y})$

$f(\boldsymbol{\lambda} \bar{x}) = \boldsymbol{\lambda} f(\bar{x})$

untuk setiap $\bar{x}$ dan $\bar{y}$ di E dan $\boldsymbol{\lambda}$ di R, sehingga f adalah anggota dari L(E,R).

Ruang vektor L(E,R) dinamakan ruang dual dari E dan diberi lambang E*. Berdasarkan definisi, maka ruang dual dari E, yakni E* adalah himpunan semua bentuk linear yang didefinisikan pada E. Sebagaimana halnya E, maka E* pun memiliki basis yang disebut basis dual (Djauhari, 1988).

2.8 Pemetaan Linear dalam Bentuk Matriks

Misalkan $\{\bar{e}_1, \bar{e}_2, \ldots, \bar{e}_n\}$ dan $\{\bar{f}_1, \bar{f}_2, \ldots, \bar{f}_p\}$ berturut-turut menyatakan basis di ruang vektor E dan F. Bila f : E → F untuk setiap j = 1,2,…,n dan i = 1,2,…,p

$$f(\bar{e}_j) = a_{1j}\,\bar{f}_1 + a_{2j}\,\bar{f}_2 + \ldots + a_{pj}\,\bar{f}_p$$

maka matriks :

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ .. & .. & \cdots & .. \\ a_{p1} & a_{p2} & \cdots & a_{pn} \end{bmatrix}$$

dinamakan matriks pemetaan dari f. Matriks A pada definisi ini berukuran p baris (dimensi F) dan n kolom (dimensi E). Jadi ditulis A(pxn). Sering dituliskan **A** = (a_{ij}) dimana a_{ij} adalah elemen baris ke-i dan kolom ke-j dari **A** (Djauhari, 1988).

2.9 Nilai dan Vektor Karakteristik

Jika **Q** adalah vektor berukuran n x 1 dan **P** adalah matriks berukuran n x n, maka **PQ** didefinisikan sebagai kombinasi linear dari unsur-unsur **Q**. Hal ini sering dikatakan sebagai transformasi linear dari **Q** ke **PQ**.

Transformasi khusus, yaitu transformasi dari **Q** ke kelipatan **Q,** yaitu yang memenuhi hubungan $\mathbf{PQ} = \boldsymbol{\lambda}\mathbf{Q}$, dengan λ adalah skalar. Hubungan tersebut terpenuhi jika dan hanya jika determinan λ I – **P** berderajat n dalam λ, sehingga terdapat n nilai λ yang memenuhi

$$| \mathbf{P} - \lambda I_n | = 0 \quad \text{atau} \quad | \lambda I_n - \mathbf{P} | = 0$$

Nilai λ ini disebut nilai eigen atau nilai karakteristik dari matriks **P** dan dilambangkan dengan $\lambda_1, \lambda_2, \ldots, \lambda_n$. Untuk setiap nilai λ_i terdapat vektor Q_i yang berukuran n x 1 yang memenuhi hubungan :

$$\mathbf{P}\, Q_i = \lambda_i\, Q_i$$

Q_i dinamakan vektor eigen atau vektor karakteristik ke-i dari **P** yang berpadanan dengan nilai λ_i.

$(\lambda_1, Q_1), (\lambda_2, Q_2), \ldots, (\lambda_n, Q_n)$ merupakan pasangan nilai karakteristik dan vektor karakteristik (Mattjik dan Sumertajaya, 2011).

2.10 Bentuk Bilinear

Misalkan E suatu ruang vektor, dim (E) = n, dan f pemetaan dari E x E ke dalam R. Pemetaan f dinamakan bentuk bilinear pada E, bila :

1). $f(\bar{x} + \bar{y}, \bar{z}) = f(\bar{x}, \bar{z}) + f(\bar{y}, \bar{z})$

2). f ($\bar{x}, \bar{y} + \bar{z}$) = f ($\bar{x}, \bar{y}$) + f ($\bar{x}, \bar{z}$)

3). f (α $\bar{x}$, β $\bar{y}$) = α β f($\bar{x}, \bar{y}$)

Untuk setiap $\bar{x}, \bar{y}, \bar{z}$ di E, dan α, β real.

Bentuk bilinear f dikatakan simetris bila f ($\bar{x}, \bar{y}$) = f ($\bar{y}, \bar{x}$) untuk setiap $\bar{x}$ dan $\bar{y}$ di E (Djauhari, 1988).

2.11 Ruang Euclides

Ruang vektor E, dengan dim (E) = n, yang dilengkapi suatu produk skalar f dinamakan ruang Euclides dan f ($\bar{x}, \bar{y}$) disebut produk skalar dari $\bar{x}$ dan $\bar{y}$. Dalam suatu ruang Euclides, dapat diukur panjang (vektor), jarak (antara dua vektor), serta sudut (yang dibentuk oleh dua vektor). Untuk itu akan dianggap E ruang Euclides dengan produk skalar f (Djauhari, 1988).

2.12 Diagram Dual

Misalkan **X**(pxn) adalah matriks data hasil pengukuran p buah variabel kuantitatif pada n individu,

$$\mathbf{X} = \begin{bmatrix} x_1^1 & x_2^1 & \dots & x_n^1 \\ x_1^2 & x_2^2 & \dots & x_n^2 \\ . & . & & . \\ . & . & & . \\ x_1^p & x_2^p & \dots & x_n^p \end{bmatrix}$$

x_i^j yang merupakan elemen baris ke-j dan kolom ke-i, adalah nilai pengukuran variabel ke-j pada individu ke-i, dimana i di I = {1,2,…,n} dan J = {1,2,…,p}.

Urutan bilangan $(x_i^1, x_i^2, \ldots, x_i^p)$ yakni urutan nilai pengukuran variabel ke-1 sampai dengan variabel ke-p pada individu ke-i, dapat dinyatakan dengan vektor :

$$\bar{x}_i = \begin{bmatrix} x_i^1 \\ x_i^2 \\ . \\ . \\ x_i^p \end{bmatrix} = \sum_{k=1}^{p} x_i^k \, \bar{e}_k$$

di E = R^p ; artinya dengan E merupakan ruang vektor real berdimensi p.

Dalam hal ini $\{\bar{e}_1, \bar{e}_2, \ldots, \bar{e}_p\}$ menyatakan basis kanonik dari ruang vektor individu E. Jadi $\bar{x}_i$ menggambarkan vektor individu ke-i di E ; i=1,2,…,n.

Demikian pula untuk urutan bilangan $(x_1^j, x_2^j, \ldots, x_n^j)$, yang merupakan urutan hasil pengukuran variabel ke-j pada individu ke-1 sampai dengan ke-n, dapat dinyatakan sebagai vektor :

$$\bar{x}^j = \begin{bmatrix} x_1^j \\ x_2^j \\ . \\ . \\ x_n^j \end{bmatrix} = \sum_{k=1}^{n} x_k^j \, \bar{f}_k$$

di F = R^n

Disini $\{\bar{f}_1, \bar{f}_2, \ldots, \bar{f}_n\}$ adalah basis kanonik dari ruang vektor variabel F. Dengan demikian, $\bar{x}^j$ menyatakan vektor variabel ke-j di F ; j = 1,2,…,p.

Cara pandang tersebut, mengakibatkan bahwa di E memiliki titik-titik individu $\{\bar{x}_i$; i=1,2,…,n$\}$ dan di F memiliki titik variabel $\{\bar{x}^j$; j=1,2,…,p$\}$. Sekarang pandang E* dan F* yakni ruang dual dari E dan dari F.

Misalkan $\{\bar{e}_1^*, \bar{e}_2^*, \ldots, \bar{e}_p^*\}$ dan $\{\bar{f}_1^*, \bar{f}_2^*, \ldots, \bar{f}_n^*\}$ adalah basis-basis dualnya.

Berdasarkan definisi basis dual, maka diperoleh :

1). $\bar{e}_j^* \ (\bar{x}_i) = \bar{e}_j^* \ (\sum_{k=1}^{p} x_i^k \ \bar{e}_k \ ) = \ < \bar{e}_j^*, \bar{x}_i > \ = x_i^j$

2). $\bar{f}_i^* \ (\bar{x}^j) = \bar{f}_i^* \ (\sum_{k=1}^{p} x_k^j \ \bar{f}_k \ ) = \ < \bar{f}_i^*, \bar{x}^j > \ = x_i^j$

Dari uraian 1) dan 2), dapat disimpulkan bahwa x_i^j adalah :

a). nilai $\bar{e}_j^*$ pada vektor individu ke-i atau dengan kata lain $\bar{e}_j^*$ menggambarkan variabel ke-j di E*.

b). $\bar{f}_i^*$ pada vektor variabel ke-j. Sehingga $\bar{f}_i^*$ menyatakan individu ke-i di F*.

Sampai disini, telah diperoleh representasi individu dan variabel sebagai berikut.

a). Terhadap variabel-j dapat dikaitkan vektor $\bar{x}^j$ di F dan bentuk linear $\bar{e}_j^*$ di E*

b). Terhadap individu-i dapat dikaitkan vektor $\bar{x}_i$ di E dan bentuk linear $\bar{f}_i^*$ di F*

Oleh karena itu dalam usaha menggambarkan mekanisme yang ada dalam induk dasar **X**, adalah wajar dipandang sebagai pemetaan.

$Q : F^* \longrightarrow E$

$\bar{f}_i^* \longrightarrow \bar{x}_i = Q(\bar{f}_i^*)$

Jadi Q memetakan bentuk linear $\bar{f}_i^*$ menjadi vektor individu $\bar{x}_i$.

Selanjutnya dapat diketahui pula bahwa matriks pemetaan dari Q tidak lain adalah matriks data **X** (pxn) itu sendiri. Dengan demikian, matriks data **X** dapat dipandang sebagai pemetaan :

$$\mathbf{X} : F^* \longrightarrow E$$

$$\bar{f}_i^* \longrightarrow \bar{x}_i = X\ \bar{f}_i^*$$

Dan $\mathbf{X}^\mathbf{T}$ transpose dari **X**, adalah pemetaan :

$$\mathbf{X}^\mathbf{T} : E^* \longrightarrow F$$

$$\bar{e}_j^* \longrightarrow \bar{x}^j = X\ \bar{e}_j^*$$

Sehingga $\mathbf{X}^\mathbf{T}$ yang memetakan bentuk linear $\bar{e}_j^*$ menjadi $\bar{x}^j$

Menyajikan himpunan individu I dan himpunan variabel J adalah salah satu tujuan utama dari teknik-teknik analisis data multidimensi. Untuk itu perlu membuat partisi dari I yang terdiri atas kelas-kelas dari individu-individu yang saling berdekatan. Jadi perlu dipertegas pengertian kedekatan antar individu. Demikian pula untuk membuat kelas-kelas dari J, perlu didefinisikan kedekatan antar variabel. Dalam analisis linear, kedekatan tersebut akan diukur dengan bantuan metrik Euclides (Djauhari, 1988).

2.12.1 Metrik

Diberikan sebarang himpunan X. Fungsi d : X x X $\longrightarrow$ R dengan R himpunan bilangan real, yang memenuhi sifat-sifat :

1. $d(x,y) \geq 0$ untuk setiap $x,y \in X$ dan $d(x,y) = 0$ jika dan hanya jika $x = y$
2. $d(x,y) = d(y,x)$ untuk setiap $x,y \in X$
3. $d(x,y) \leq d(x,z) + d(y,z)$ untuk setiap $x,y,z \in X$

disebut metrik atau jarak pada X. Himpunan X yang dilengkapi dengan suatu metrik d ditulis dengan (X,d) disebut ruang metrik. Anggota ruang metrik (X,d) disebut titik dan untuk setiap x,y ϵ X bilangan non-negatif, d(x,y) disebut jarak titik x ke titik y.

2.12.2 Norm

Definisi Norm :

Misalkan P adalah suatu ruang vektor atas R ruang bilangan real. Suatu fungsi : $||.|| : P \longrightarrow R$ disebut norm vektor jika untuk semua x,y ϵ P berlaku :

- $||\, x \,|| \geq 0$
- $||\, x \,|| = 0$ jika dan hanya jika x = 0
- $||\, c\, x \,|| = |\, c \,|\ ||\, x \,||$ untuk setiap c anggota himpunan bilangan real
- $||\, x + y \,|| \leq ||\, x \,|| + ||\, y \,||$

2.12.3 Metrik di E

Misalkan E ruang Euclides dengan metrik M yang berperan mengukur kedekatan antar individu. Dengan memandang M sebagai isomorfisma (sebuah fungsi yang bersifat bijektif) dari E pada E* maka sangatlah wajar bila pada F* diterapkan metrik W sedemikian sehingga :

$$||\, \bar{x}_i - \bar{x}_k \,||_M = ||\, \bar{f}_i^* - \bar{f}_k^* \,||_W$$

dimana X $(\bar{f}_i^*)$ = $\bar{x}_i$; i=1,2,…,n. Mekanisme di atas, dapat disajikan dalam diagram berikut.

$$\begin{tikzcd} E & F^* \arrow[l, "X"'] \arrow[d, "W"] \\ E^* \arrow[u, leftarrow, "M"] \arrow[r, "X^T"'] & F \end{tikzcd}$$

Secara umum, untuk setiap $\bar{a}$ dan $\bar{b}$ di F* diinginkan W

$\| X(\bar{a}) - X(\bar{b}) \|_M = \| \bar{a} - \bar{b} \|_W$

Ini berarti pula bahwa untuk setiap $\bar{a}$ di F* berlaku :

$\| X(\bar{a}) \|_M = \| \bar{a} \|_W$ (Djauhari, 1988).

2.12.4 Metrik di F

Seperti halnya di E, maka untuk mengukur kedekatan antar variabel serta melihat kolinearitasnya di F diterapkan suatu metrik di N. Dengan mengikuti konsep yang sama seperti untuk E, di E* pun dapat diterapkan metrik V $\ni$

$\| \bar{x}^j - \bar{x}^k \|_N = \| \bar{e}_j^* - \bar{e}_k^* \|_V$

Ingatlah bahwa $X^T(\bar{e}_j^*) = \bar{x}^j$; j = 1,2,…,p.

Secara umum, hal tersebut berarti bahwa untuk setiap $\bar{a}$ di E* berlaku :

$\| X^T(\bar{a}) \|_N = \| \bar{a} \|_V$

Hubungan antara V dan N diberikan pada dalil berikut.

Dalil 2.10.1.2 :

Jika untuk setiap $\bar{a}$ di E* berlaku $\| X^T(\bar{a}) \|_N = \| \bar{a} \|_V$, maka diperoleh diagram berikut :

$$
\begin{tikzcd}
E & F^* \arrow[l, "X"'] \\
E^* \arrow[u, "V"] \arrow[r, "X^T"'] & F \arrow[u, "N"']
\end{tikzcd}
$$

Komutatif, artinya : **V = X N X^T^** $\mathbf{V = X\,N\,X^T}$

Selanjutnya kedua diagram di atas dapat digabungkan sehingga membentuk diagram berikut ini yang menggambarkan seluruh mekanisme dasar yang ada bila dihadapkan dengan matriks data **X (pxn)**.

$$
\begin{tikzcd}
R^P = E \arrow[d, shift right, "M"'] & F^* \arrow[l, "X"'] \arrow[d, shift right, "W"'] \\
E^* \arrow[u, shift right, "V"'] \arrow[r, "X^T"'] & F = R^n \arrow[u, shift right, "N"']
\end{tikzcd}
$$

Diagram tersebut selanjutnya disebut diagram dual (Djauhari, 1988).

2.12.5 Metrik Bobot di F

Bila X (pxn) matriks data yang terdiri dari p baris dan n kolom maka di E = R^P dengan himpunan vektor individu { $\bar{x}_i$ / i =1,2,…,n}.

Misalkan terhadap setiap individu -i, jadi terhadap setiap vektor $\bar{x}_i$ di himpunan itu, diberikan bobot p_i ($p_i > 0$ dan $\sum_{i=1}^{n} p_i = 1$)

Dengan demikian, maka vektor $\bar{g}$ dimana :

$\bar{g} = \sum_{i=1}^{n} p_i \bar{x}_i$

tidak lain adalah vektor mean atau pusat gravitasi dari himpunan vektor individu di atas. Elemen ke-j dari $\bar{g}$ adalah

$g_j = \sum_{i=1}^{n} p_i \; x_i^j$

yang merupakan mean sampel untuk variabel ke-j. Khususnya jika $p_i = \frac{1}{n}$; untuk setiap i = 1,2,…,n yang berarti bahwa semua individu memiliki bobot yang sama (ini akan diperoleh bila sampel diambil secara acak); maka

$\bar{g} = \frac{1}{n} \sum_{i=1}^{n} x_i$ dan $g_j = \frac{1}{n} \sum_{i=1}^{n} x_i^j$

Di dalam analisis data univariate, besaran-besaran yang sering digunakan terutama adalah mean sampel (rata-rata) dan variansi sampel. Sebagai pengembangannya pada analisis data multivariate, akan digunakan vektor mean $\bar{g}$ dan momen inersia (yang merupakan perluasan dari variansi) (Djauhari, 1988).

<u>Definisi 2.12.5.1</u>

Momen inersia individu $\bar{x}_i$ yang berbobot p_i terhadap suatu titik $\bar{a}$ di E adalah :

$p_i \; ||\bar{x}_i - \bar{a}||_M^2$

(bobot dikalikan dengan kuadrat jarak).

<u>Definisi 2.12.5.2</u>

Momen inersia himpunan vektor individu { $\bar{x}_i$ | i=1,2,…,n}, dimana $\bar{x}_i$ berbobot p_i, terhadap suatu titik $\bar{a}$ di E adalah :

$I_{\bar{a}} = \sum_{i=1}^{n} p_i \, || \, \bar{x}_i - \bar{a} \, || \,_M$

Dalil 2.12.5.1 :

Untuk setiap $\bar{a}$ di E, berlaku :

$$I_{\bar{a}} = I_{\bar{g}} + ||\bar{g} - \bar{a}||_M^2$$

Bukti.

Karena $\bar{x}_i$ - $\bar{a}$ = ($\bar{x}_i$ - $\bar{g}$) + ($\bar{g}$ - $\bar{a}$), maka

$$||\bar{x}_i - \bar{a}||_M^2 = ||\bar{x}_i - \bar{g}||_M^2 + ||\bar{g} - \bar{a}||_M^2 + 2M(\bar{x}_i - \bar{g}, \bar{g} - \bar{a})$$

Sedangkan,

$$\sum_{i=1}^{n} p_i \, M(\bar{x}_i - \bar{g}, \bar{g} - \bar{a}) = M\left(\sum_{i=1}^{n} p_i (\bar{x}_i - \bar{g}), \bar{g} - \bar{a}\right)$$

$$= M\left(\sum_{i=1}^{n} p_i \; \bar{x}_i \; - \; \bar{g} \sum_{i=1}^{n} p_i, \bar{g} - \bar{a}\right)$$

$= M(\bar{g} - \bar{g}, \bar{g} - \bar{a})$, sebab $\sum_{i=1}^{n} p_i \; \bar{x}_i = \bar{g}$ dan $\sum_{i=1}^{n} p_i = 1$

$$= M(\bar{0}, \bar{g} - \bar{a}) = 0$$

Jadi :

$$I_{\bar{a}} = \sum p_i ||\bar{x}_i - \bar{a}||_M^2$$

$$= \sum p_i \{ ||\bar{x}_i - \bar{g}||_M^2 + ||\bar{g} - \bar{a}||_M^2 \}$$

$$= \sum p_i ||\bar{x}_i - \bar{g}||_M^2 + ||\bar{g} - \bar{a}||_M^2 \sum_{i=1}^{n} p_i$$

$$= I_{\bar{g}} + ||\bar{g} - \bar{a}||_M^2.$$

Dalil ini dikenal dengan nama Dalil Huyghens, yang menyimpulkan bahwa : Vektor mean $\bar{g}$ adalah vektor yang meminimumkan $I_{\bar{a}}$. Artinya, $I_{\bar{a}}$ akan minimum jika dan hanya jika $\bar{a} = \bar{g}$.

Karena translasi titik $\bar{0}$ di E ke titik $\bar{g}$ tidak berpengaruh terhadap bentuk atau konfigurasi himpunan vektor individu, untuk selanjutnya dianggap $\bar{g} = \bar{0}$. Ini berarti bahwa $g_j = 0$ untuk setiap j = 1, 2, …, p. Dengan kata lain, untuk selanjutnya dianggap variabel $\bar{x}^j$ terpusat, artinya memiliki mean = 0 atau g_j = 0 untuk setiap j = 1, 2, …, p. Akibatnya, untuk setiap j = 1, 2, …, p berlaku :

1). Variansi dari $\bar{x}^j$ adalah :

$$\begin{aligned}\mathrm{Var}\,(\bar{x}^j) &= \sum p_i \,(x_i^j - g_j)^2\\ &= \sum p_i \;(x_i^j)^2 - (g_j)^2\\ &= \textstyle\sum_{i=1}^n p_i \;(x_i^j)^2\end{aligned}$$

2). Kovariansi antara variabel $\bar{x}^j$ dan $\bar{x}^k$ adalah :

$$\begin{aligned}\mathrm{Cov}\,(\bar{x}^j, \bar{x}^k) &= \textstyle\sum_{i=1}^n p_i \,(x_i^j - g_j)\,(x_i^k - g_k)\\ &= \textstyle\sum_{i=1}^n p_i \, x_i^j \, x_i^k - g_j \, g_k = \sum_{i=1}^n p_i \, x_i^j \, x_i^k\end{aligned}$$

Variansi dan kovariansi di atas dapat dituliskan dalam bentuk perkalian matriks sebagai berikut.

Var $(\bar{x}^j) = (\bar{x}^j)^T D_P \, \bar{x}^j$ dan Cov $(\bar{x}^j, \bar{x}^k) = (\bar{x}^j)^T D_P \, \bar{x}^k$ dimana

$$D_P = \begin{bmatrix} p_1 & 0 & \dots & 0 \\ 0 & p_2 & \dots & 0 \\ . & . & & . \\ . & . & & . \\ 0 & 0 & \dots & p_n \end{bmatrix}$$

yakni matriks diagonal ; elemen diagonal ke-i adalah bobot individu ke-i atau p_i. Oleh karena itu pada ruang variabel F akan diterapkan metrik Euclides D_P, yang dikenal dengan metrik bobot. Dengan metrik ini, maka baris { $\bar{f}_1, \bar{f}_2, \dots, \bar{f}_n$ } dari F memiliki sifat :

$D_P\,(\bar{f}_i, \bar{f}_k) = 0$ bila $i \neq k$, dan $D_P\,(\bar{f}_i, \bar{f}_i) = \|\,\bar{f}_i\,\|_{D_P}^2 = p_i$.

Dari uraian di atas, pada ruang variabel F dengan metrik D_P, diperoleh :

1). Cov $(\bar{x}^j, \bar{x}^k) = D_P\ (\bar{x}^j, \bar{x}^k)$. Artinya, kovariansi tidak lain adalah produk skalar.

2). Var $(\bar{x}^j) = D_P\ (\bar{x}^j, \bar{x}^j) = \ ||\ \bar{x}^j\ ||^2_{D_P}$. Jadi, variansi adalah kuadrat norm. Sedangkan norm tidak lain adalah standar deviasi.

Sebagai akibat dari pemberian metrik D_P pada F, maka sudut 2 vektor di F akan berkaitan dengan korelasi. Hal ini akan dirumuskan pada dalil berikut.

Dalil 2.12.5.2 :

Misalkan θ adalah sudut antara $\bar{x}^j$ dan $\bar{x}^k$; j, k = 1, 2, …, p. Maka cos θ sama dengan koefisien korelasi antara x^j dan x^k.

Bukti.

$$\text{Cos } \theta = \frac{D_P\ (\bar{x}^j, \bar{x}^k)}{||\ \bar{x}^j\ ||_{D_P}\ ||\ \bar{x}^k\ ||_{D_P}} = \frac{\text{Cov } (\bar{x}^j, \bar{x}^k)}{\sqrt{Var\ (\bar{x}^j)}\ \sqrt{Var\ (\bar{x}^k)}} = \text{r } (\bar{x}^j, \bar{x}^k)$$

dimana r $(\bar{x}^j, \bar{x}^k)$ adalah koefisien korelasi antara $\bar{x}^j$ dan $\bar{x}^k$.

Selanjutnya, dengan memandang diagram dual :

$$\begin{tikzcd}
R^P = E \arrow[d, shift right, "M"'] & F^* \arrow[l, "X"'] \arrow[d, shift right, "W"'] \\
E^* \arrow[u, shift right, "V"'] \arrow[r, "X^T"'] & F = R^n \arrow[u, shift right, "D_P"']
\end{tikzcd}$$

diperoleh pula dalil berikut.

Dalil 2.12.5.3 :

Dengan menganggap bahwa X terpusat (artinya variabel $\bar{x}^j$ memiliki mean sama dengan 0 untuk setiap j = 1, 2, …, p) maka matriks **V = X $\boldsymbol{D_P}$ X$^{\mathbf{T}}$** adalah matriks variansi-kovariansi. Dengan kata lain, elemen baris ke-j dan kolom ke-k dari **V** adalah $V_{jk} =$ Cov $(\bar{x}^j, \bar{x}^k)$ (Djauhari, 1988).

Bukti.

$$\mathbf{V} = \mathbf{X}\,\boldsymbol{D_P}\,\mathbf{X^T} = \begin{bmatrix} x_1^1 & x_2^1 & \dots & x_n^1 \\ x_1^2 & x_2^2 & \dots & x_n^2 \\ . & . & & . \\ . & . & & . \\ x_1^p & x_2^p & \dots & x_n^p \end{bmatrix} \begin{bmatrix} p_1 & 0 & \dots & 0 \\ 0 & p_2 & \dots & 0 \\ . & . & & . \\ . & . & & . \\ 0 & 0 & \dots & p_n \end{bmatrix} \begin{bmatrix} x_1^1 & x_1^2 & \dots & x_1^p \\ x_2^1 & x_2^2 & \dots & x_2^p \\ . & . & & . \\ . & . & & . \\ x_n^p & x_2^p & \dots & x_n^p \end{bmatrix}$$

$$= \begin{bmatrix} x_1^1 & x_2^1 & \dots & x_n^1 \\ x_1^2 & x_2^2 & \dots & x_n^2 \\ . & . & & . \\ . & . & & . \\ x_1^p & x_2^p & \dots & x_n^p \end{bmatrix} \begin{bmatrix} p_1x_1^1 & p_1x_1^2 & \dots & p_1x_1^p \\ p_2x_2^1 & p_2x_2^2 & \dots & p_2x_2^p \\ .. & .. & \dots & .. \\ .. & .. & \dots & .. \\ p_nx_n^1 & p_nx_n^2 & \dots & p_nx_n^p \end{bmatrix}$$

$$V_{jk} = \begin{pmatrix} x_1^j & x_2^j & \dots & x_n^j \end{pmatrix} \begin{bmatrix} p_1x_1^k \\ p_2x_2^k \\ .. \\ .. \\ p_nx_n^k \end{bmatrix}$$

$$= \sum_{i=1}^{n} p_i\, x_i^j\, x_i^k$$

$= $ Cov $(\bar{x}^j, \bar{x}^k)$.

Catatan :

Elemen diagonal ke-j dari V adalah Var $(\bar{x}^j)$

Akibat.

Misalkan metrik di E adalah :

$$M = D_{1/\sigma^2} = \begin{bmatrix} \frac{1}{\sigma_1^2} & 0 & \dots & 0 \\ 0 & \frac{1}{\sigma_2^2} & \dots & 0 \\ . & . & & . \\ . & . & & . \\ 0 & 0 & \dots & \frac{1}{\sigma_p^2} \end{bmatrix}$$

yakni metrik diagonal ; elemen diagonal ke-j adalah $\frac{1}{\sigma_j^2}$ dan σ_j^2 = var $(\bar{x}^j)$ atau variansi dari variabel -$\bar{x}^j$. Dengan menuliskan M = $D_{1/\sigma^2}{}^T \ D_{1/\sigma^2}$, dimana :

$$D_{1/\sigma} = \begin{bmatrix} \frac{1}{\sigma_1} & 0 & \dots & 0 \\ 0 & \frac{1}{\sigma_2} & \dots & 0 \\ . & . & & . \\ . & . & & . \\ 0 & 0 & \dots & \frac{1}{\sigma_p} \end{bmatrix}$$

maka diperoleh diagram dual sebagai berikut :

$$\begin{tikzcd}
R^P = E \arrow[d, shift right, "I_p"'] & E \arrow[l, "D_{1/\sigma}"'] \arrow[d, shift right, "M"'] & F^* \arrow[l, "X"'] \arrow[d, shift right, "W"'] \\
E^* \arrow[u, shift right, "R"'] \arrow[r, "{D_{1/\sigma}}^T"'] & E^* \arrow[u, shift right, "V"'] \arrow[r, "X^T"'] & F = R^n \arrow[u, shift right, "D_p"']
\end{tikzcd}$$

Komutatif dan :

1). Transformasi matriks data X menjadi DX mengakibatkan bahwa metrik yang digunakan di E tidak lagi M tapi metrik klasik I_p.

2). Matriks R = $D_{1/\sigma}$ X D_p X^T ${D_{1/\sigma}}^T$ tidak lain adalah matriks korelasi. Semua elemen diagonal dari R berharga 1. Elemen baris ke-j dan kolom ke-k dari R adalah r ($\bar{x}^j, \bar{x}^k$) yakni koefisien korelasi antara variabel $\bar{x}^j$ dan variabel $\bar{x}^k$ (Djauhari, 1988).

Bila matriks data **X** terpusat, maka diketahui :

a). Norm di F sama dengan standar deviasi

b). Produk skalar di F ekuivalen dengan kovariansi

c). Cosinus sudut di F adalah koefisien korelasi

d). Titik nol dari E berada di $\bar{g} = \bar{o}$

e). V = **X** D_p **X’** merupakan matriks variansi kovariansi (Djauhari, 1988).

2.13 Analisis Peubah Ganda

Misalkan $\boldsymbol{X_i}$ adalah vektor acak berukuran p x 1 atau

$$\boldsymbol{X_i} = \begin{bmatrix} x_i^1 \\ x_i^2 \\ . \\ . \\ x_i^p \end{bmatrix}$$

maka setiap $\boldsymbol{X_i}$ adalah vektor acak dan diasumsikan $x_i^1, x_i^2, \ldots, x_i^p$ mungkin saling tidak bebas. Nilai tengah dari vektor acak $\boldsymbol{X_i}$ dinotasikan **μ** dan matriks kovarian dari $\boldsymbol{X_i}$ dinotasikan **Σ** didefinisikan sebagai :

$$\text{E}(\mathbf{X}) = \begin{bmatrix} E(x_1) \\ E(x_2) \\ . \\ . \\ E(x_p) \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ . \\ . \\ \mu_p \end{bmatrix} = \boldsymbol{\mu}$$

$$\boldsymbol{\Sigma} = \begin{bmatrix} \text{Var}(x_1) & \text{Cov}(x_1, x_2) & \ldots & \text{Cov}(x_1, x_p) \\ \text{Cov}(x_2, x_1) & \text{Var}(x_2) & \ldots & \text{Cov}(x_2, x_p) \\ .. & .. & \ldots & .. \\ .. & .. & \ldots & .. \\ \text{Cov}(x_p, x_1) & \text{Cov}(x_p, x_2) & \ldots & \text{Var}(x_p) \end{bmatrix} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & .. & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & .. & \sigma_{2p} \\ .. & .. & .. & .. \\ .. & .. & .. & .. \\ \sigma_{p1} & \sigma_{p1} & .. & \sigma_{pp} \end{bmatrix}$$

Dengan mengartikan Cov (x_i, x_i) = Var $(x_i) = \sigma_{ii}$, bentuk Cov (x_i, x_j) disebut sebagai unsur ke (i, j) dari matriks **Σ**, dimana :

σ_{ii} = Var (x_i) = E $[(x_i - \mu_i)^2]$ untuk i = 1, 2, …, p

σ_{ij} = Cov (x_i, x_j) = E $[(x_i$ - $\mu_i)(x_j$ - $\mu_j)]$ untuk i ≠ j = 1, 2, …, p (Johnson, 1998).

2.14 Analisis Komponen Utama

Analisis Komponen Utama (AKU) biasanya digunakan untuk mendapatkan komponen utama dengan mempertahankan sebanyak mungkin keragaman dalam data, tanpa menghilangkan sebagian besar informasi yang terkandung pada data asal. Komponen utama cukup dapat memberikan informasi yang besar seperti yang diterangkan oleh p-peubah awal. Sehingga sejumlah komponen utama tersebut dapat menggantikan sebanyak p-peubah, artinya data awal yang terdiri dari n-individu dengan p-peubah kemudian direduksi menjadi n-individu dengan beberapa komponen utama.

Melalui pendekatan ruang dual, komponen – komponen utama ini akan diperoleh dengan melibatkan nilai karakteristik dan vektor karakteristik dari matriks hasil kali bentuk bilinear **W** pada F* dengan matriks diagonal bobot $\boldsymbol{D_p}$. Misalkan $\bar{C}^i$ adalah komponen utama ke-i dengan i=1,2,...k, komponen utama $\bar{C}^i$ dapat ditentukan yaitu $\bar{C}^i = \alpha_i \, \bar{a}_i$ dengan $\bar{a}_i$ adalah vektor karakteristik dan α_i adalah suatu skalar yang diperoleh dari persamaan berikut :

$(p_i (\alpha_i^2)) V_i' V_i = \lambda_i$ dengan p_i adalah bobot individu ke- i

Bentuk bilinear **W** seperti pada uraian di atas, dapat ditentukan melalui matriks data terpusat yaitu selisih dari data sebenarnya dengan rata-ratanya. Sehingga komponen utama $\bar{C}^i$ yang dihasilkan dengan i=1,2,...k, bersifat terpusat pula (Djauhari, 1988).

III. METODOLOGI PENELITIAN

3.1 Waktu dan Tempat Penelitian

Penelitian ini dilakukan pada Semester Ganjil Tahun Akademik 2015/2016, bertempat di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.

3.2 Metode Penelitian

Penulisan skripsi ini dilakukan dengan menggunakan studi literatur secara sistematis yang diperoleh dari buku-buku atau media lain untuk mendapatkan informasi sebanyak mungkin untuk mendukung penulisan skripsi ini, kemudian melakukan simulasi sebagai aplikasi untuk menjelaskan teori yang telah didapat.

Adapun langkah-langkah yang dilakukan dalam penelitian ini adalah sebagai berikut :

1. Mempelajari konsep-konsep matriks yang dalam hal ini nantinya akan digunakan untuk mempermudah dalam penyajian data multivariate.
2. Mendefinisikan variabel pertama, kedua, dan seterusnya serta menentukan ada berapa pengamatan pada setiap variabel tersebut.

3. Memberikan matriks data untuk data *multivariate* yang telah didefinisikan sebelumnya.
4. Menentukan vektor mean $\bar{g}$ dengan elemen-elemennya merupakan mean dari setiap variabel yang diamati.
5. Menentukan matriks data terpusat dari matriks data mentah. Dalam hal ini untuk memperoleh matriks data terpusat **Y** dari matriks data mentah $\mathbf{X}_{(pxn)}$ dapat dilakukan dalam bentuk operasi matriks **Y = X – G** ; dengan semua vektor kolom dari $\mathbf{G}_{(pxn)}$ sama dengan $\bar{g}$.
6. Dari matriks data terpusat yang dihasilkan pada langkah ke-5, akan ditentukan bentuk bilinear **W = Y' M Y** dengan **M** = I_p dan bobot semua individu sama.
7. Dengan bobot semua individu sama dan p_i menjelaskan bobot individu ke-i maka diketahui $\boldsymbol{D_p}$ = $p_i\ I_n = \frac{1}{n}\ I_n$; dengan i=1,2,...,n. Sehingga selanjutnya akan ditentukan matriks hasil kali dari **W** dengan $\boldsymbol{D_p}$.
8. Setelah diperoleh matriks **W** $\boldsymbol{D_p}$, maka akan ditentukan $\boldsymbol{\lambda}$ yaitu nilai-nilai karakteristik dan V yaitu vektor karakteristiknya.
9. Menentukan komponen-komponen utamanya yaitu $\bar{C}^1$, $\bar{C}^2$, dan seterusnya dengan $\bar{C}^i = \alpha_i\ V_i$; dengan i = 1,2,...n . V_i adalah vektor karakteristik ke- i dan α_i adalah suatu skalar yang diperoleh dari persamaan berikut :

$$(p_i\ (\alpha_i^2))\ V_i'\ V_i = \lambda_i$$

10. Menentukan persentase kualitas global yang dimiliki setiap komponen utama yang dihasilkan yaitu $Q_i = \frac{\lambda_i}{Tr\ (W\ D_p\ )}$; dengan Q_i adalah persentase kualitas global komponen utama ke- i.

V. KESIMPULAN

Dari pembahasan dan analisis yang telah dilakukan, maka diperoleh kesimpulan sebagai berikut :

1. Melalui pendekatan ruang dual, analisis komponen utama berusaha mereduksi dimensi ruang. Pada penelitian ini analisis komponen utama berhasil mereduksi dimensi ruang variabel (E = R^p), yang tadinya p menjadi q dengan q < p. Dalam proses mereduksi dimensi ruang melalui pendekatan ruang dual digunakan bentuk bilinear dari matriks data terpusat, sehingga selanjutnya analisis diteruskan hanya dengan melibatkan bentuk bilinear itu saja tanpa melibatkan lagi matriks data awal.
2. Komponen utama yang diperoleh merupakan kombinasi linear dari peubah-peubah yang diamati, informasi yang terkandung pada komponen utama merupakan gabungan dari semua peubah dengan bobot tertentu. Kombinasi linear yang dipilih merupakan kombinasi linear dengan ragam paling besar yang memuat informasi paling banyak. Seberapa besar keragaman tersebut dapat dijelaskan, diketahui dengan mencari kualitas globalnya.

DAFTAR PUSTAKA

Anton, H. 2000. *Aljabar Linear Elementer*. Ed. ke-7. Jilid 1. Interaksa, Batam Center.

Djauhari, M.A. 1988. *Struktur Data Statistik*. Karunika Universitas Terbuka, Jakarta.

Gaspersz, V. 1995. *Teknik Analisis dalam Penelitian Percobaan 2*. Tarsito, Bandung.

Johnson, R.A. and Wichern, D.W. 1998. *Applied Multivariate Statistical Analysis*. Prentice Hall Englewood, New Jersey.

Mattjik, A.A. dan Sumertajaya, I.M. 2011. *Sidik Peubah Ganda*. IPB PRESS, Bogor.