

# Penggunaan *Kernel Principal Component Analysis* Fungsi Polinomial Dalam Menyelesaikan Masalah Pengelompokan Plot Peubah Ganda

Sueharti Maatuil<sup>1</sup>, Hanny A. H. Komalig<sup>2</sup>, Charles Mongi<sup>3</sup>

<sup>1</sup>Program Studi Matematika, FMIPA, Universitas Sam Ratulangi, [suehartim@gmail.com](mailto:suehartim@gmail.com)

<sup>2</sup>Program Studi Matematika, FMIPA, Universitas Sam Ratulangi, [Hanyoy07@yahoo.com](mailto:Hanyoy07@yahoo.com)

<sup>3</sup>Program Studi Matematika, FMIPA, Universitas Sam Ratulangi, [charlesmongi@ymail.com](mailto:charlesmongi@ymail.com)

## Abstrak

Tujuan dari penelitian ini yaitu mempelajari penggunaan *kernel PCA* fungsi polinomial untuk membantu menyelesaikan masalah plot peubah ganda terutama yang berhubungan dalam pengelompokan. Data yang digunakan dalam penelitian ini adalah data sekunder yang berupa plot peubah ganda. Metode kernel adalah salah satu cara untuk mengatasi kasus-kasus yang tidak linier. *Kernel PCA* merupakan *PCA* yang diaplikasikan pada input data yang telah ditransformasikan ke *feature space*. Misalkan  $\Phi: \mathbf{R}^n \rightarrow \mathbf{F}$  fungsi yang memetakan semua input data  $x_i \in \mathbf{R}^n$ , berlaku  $\Phi(x_i) \in \mathbf{F}$ . Salah satu kernel yang banyak digunakan adalah kernel polinomial. Dimana  $h_0$  adalah parameter skala yang akan dipilih. Fungsi kernel polinomial  $K(x_i, x_j) = (x_i^T, x_j + h_0)^d$ . Hasil dari penelitian ini menunjukkan bahwa penggunaan Kernel Principal Component Analysis (KPCA) dengan fungsi kernel polinomial sangat membantu dalam menyelesaikan masalah plot peubah ganda yang belum dapat dikelompokkan dengan garis pemisah yang linier.

**Kata kunci** : Kernel PCA, Kernel PCA Fungsi Polinomial, Plot Peubah Ganda

## *The Use of Kernel Principal Component Analysis (KPCA) Polynomial Function in Solving Multivariate Plot Grouping*

### Abstract

The purpose of this research is to study the use of Kernel PCA polynomial functions to help solve problems of plot multivariate especially in the grouping. In this research the data used is secondary data such as multivariate plot. Kernel method is one way to overcome cases that are not linear. Kernel PCA as a PCA that applied to the data input that has been transformed to feature space. Let  $\Phi: \mathbf{R}^n \rightarrow \mathbf{F}$  function that maps all data input  $x_i \in \mathbf{R}^n$ , applies  $\Phi(x_i) \in \mathbf{F}$ . One of the kernel that is widely used is the polynomial kernel. Where  $h_0$  is the scale parameter will be selected. Polynomial kernel function  $K(x_i, x_j) = (x_i^T, x_j + h_0)^d$ . Results from this study showed use of Kernel Principal Component Analysis (KPCA) with a polynomial kernel function is very helpful in solving the problem of multivariate plot that cannot be grouped with the linear dividing line.

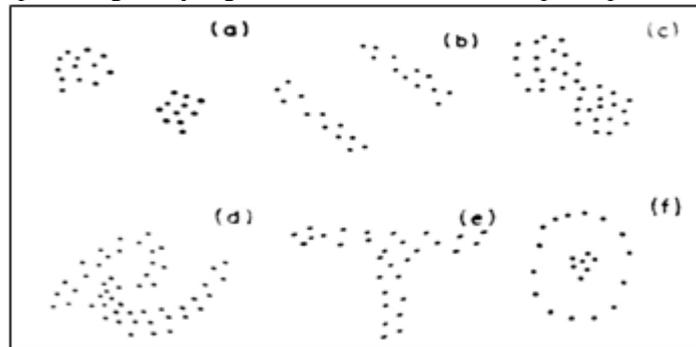
**Keywords**: Kernel PCA, Kernel PCA Polynomial Function, Plot Multivariate

## 1. Pendahuluan

Di era globalisasi sekarang ini, ilmu statistika bukan lagi hanya berupa perhitungan angka-angka tetapi telah digunakan juga dalam disiplin ilmu yang lain. Dalam suatu penelitian, analisis statistika sangat berperan penting dalam menganalisis, menginterpretasi dan mempresentasikan data. Untuk mempermudah dalam menginterpretasikan dan mempresentasikan data suatu penelitian, maka ditampilkan dalam bentuk gambar yang didalamnya menerangkan informasi yang diperlukan. Gambar berupa titik dengan posisi jarak relatif objek-objek data disebut dengan plot. Salah satu plot yang didalamnya dapat diamati pola hubungan antara dua peubah adalah *scatter plot* sedangkan untuk menampilkan skor komponen dari sebaran data ditampilkan pada *score plot*. Dalam analisis statistika, analisis *multivariate* atau analisis peubah ganda merupakan analisis yang banyak menampilkan plot dalam menyampaikan informasi dari data penelitian.

Analisis peubah ganda digunakan untuk menganalisis lebih dari satu peubah, baik peubah bebas maupun terikat pada saat yang bersamaan dan untuk menganalisa hubungan dari peubah-peubah. Salah satu Analisis peubah ganda yang digunakan untuk analisis identifikasi, prediksi, eksplorasi, deskripsi yaitu *Principal Component Analysis* yang biasa disingkat *PCA*. *PCA* digunakan untuk mereduksi dimensi data menjadi dimensi yang lebih kecil. *PCA* sebagai salah satu

teknik untuk menyederhanakan data peubah banyak yang berkorelasi tinggi dengan cara mentransformasi data secara linier sehingga diperoleh peubah baru yang dimensinya lebih kecil dengan varians maksimum dan bersifat independen, yang disebut dengan komponen utama. *PCA* hanya dapat menyelesaikan masalah linier sehingga untuk menyelesaikan masalah nonlinier digunakan *Kernel PCA* sebagai perluasan dari *PCA*. *Kernel PCA* memetakan data asal ke ruang fitur. Fungsi kernel merupakan hasil kali dalam dari pemetaan ke ruang fitur. Dengan digunakan fungsi kernel dalam dimensi yang lebih tinggi dari ruang fitur, maka kita dapat memperoleh komponen utama secara lebih efisien. *Kernel PCA* membantu dalam menyelesaikan masalah pengelompokan plot peubah ganda yang tidak linier dan belum dapat dipisahkan dengan *PCA*.

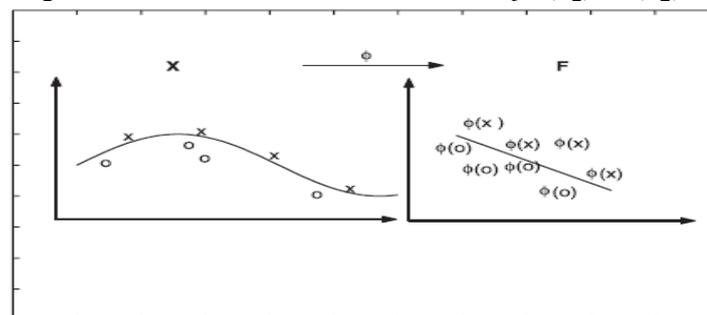


Gambar 1. Plot Peubah Ganda

Dalam menyelesaikan masalah plot peubah ganda seperti gambar plot sebaran data (d,e, f) pada Gambar 1 dapat digunakan *kernel PCA* fungsi polinomial. Fungsi kernel polinomial sebagai salah satu fungsi kernel yang sering digunakan dalam *PCA*. Pada fungsi kernel polinomial, matriks komponen utama dapat diperoleh melalui formula primal karena hasil pemetaan di ruang fitur yang jelas.

## 2. Kernel PCA

*Kernel PCA* [1] merupakan *PCA* yang diaplikasikan pada ruang input data yang telah ditransformasikan ke ruang *feature*. Misalkan  $\Phi: \mathbb{R}^n \rightarrow \mathbf{F}$  fungsi yang memetakan semua input data  $x_i \in \mathbb{R}^n$ , maka berlaku  $\Phi(x_i) \in \mathbf{F}$ . Berdasarkan transformasi ini, terlihat bahwa ruang *feature* dibangun oleh vektor-vektor  $\{\Phi(x_1), \Phi(x_2), \dots, \Phi(x_m)\}$ , sehingga semua vektor di ruang fitur dapat dinyatakan sebagai kombinasi linier dari vektor-vektor  $\{\Phi(x_1), \Phi(x_2), \dots, \Phi(x_m)\}$ .



Gambar 2. Ilustrasi pemetaan data ke ruang fitur

Sering kali fungsi  $\Phi(x)$  tidak tersedia atau tidak bisa dihitung, tetapi perkalian titik dari dua vektor dapat dihitung baik di dalam ruang input maupun di ruang fitur. Dengan kata lain, sementara  $\Phi(x)$  mungkin tidak diketahui, perkalian titik  $\langle \Phi(x_i), \Phi(x_j) \rangle$  masih bisa dihitung di ruang fitur. Suatu fungsi kernel  $k(x_i, x_j)$ , bisa untuk menggantikan perkalian titik  $\langle \Phi(x_i), \Phi(x_j) \rangle$ . Kemudian di ruang fitur, kita bisa membuat suatu garis pemisah linier yang mewakili fungsi nonlinier di ruang input. Pada Gambar 2 mendeskripsikan suatu contoh pemetaan fitur dari ruang dua dimensi ke ruang fitur dua dimensi. Dalam ruang input, data tidak bisa dipisahkan secara linier, tetapi dapat dipisahkan dengan baik di ruang fitur [2].

PCA menemukan sumbu utama dengan mendiagonalkan matriks peragam

$$C = \frac{1}{m} \sum_{j=1}^m x_j x_j^T \tag{1}$$

dimana  $v$  adalah vektor eigendan  $C$  positive definit, sehingga dapat didiagonalkan dengan nilai eigen nonnegatif

$$\lambda v = Cv \tag{2}$$

Dengan mensubstitusi persamaan (1) ke dalam persamaan (2), diperoleh

$$Cv = \frac{1}{m} \sum_{j=1}^m x_j x_j^T v \tag{3}$$

sehingga

$$v = \frac{1}{m\lambda} \sum_{j=1}^m x_j x_j^T v \tag{4}$$

$$= \frac{1}{m\lambda} \sum_{j=1}^m (x_j \cdot v) x_j$$

Dapat ditunjukkan bahwa  $(xx^T)v = (x \cdot v)x$

$$\begin{aligned} (xx^T)v &= \begin{pmatrix} x_1 x_1 & x_1 x_2 & \dots & x_1 x_M \\ x_2 x_1 & x_2 x_2 & \dots & x_2 x_M \\ \vdots & \vdots & \ddots & \vdots \\ x_M x_1 & x_M x_2 & \dots & x_M x_M \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_M \end{pmatrix} \\ &= (x_1 v_1 + x_2 v_2 + \dots + x_M v_M) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_M \end{pmatrix} \\ &= (x \cdot v)x \end{aligned} \tag{5}$$

$(x_j \cdot v)$  merupakan skalar, jadi ini berarti bahwa semua solusi  $v$  dengan  $\lambda \neq 0$  terletak pada rentang  $x_1, \dots, x_m$ , yaitu

$$v = \sum_{i=1}^m a_i x_i \tag{6}$$

Matriks peragam di ruang fitur untuk vektor  $\{\Phi(x_1), \Phi(x_2), \dots, \Phi(x_m)\}$  dapat dituliskan sebagai

$$C = \frac{1}{m} \sum_{j=1}^m \Phi(x_j) \Phi(x_j)^T \tag{7}$$

dan masalah nilai eigendi ruang fitur  $F$  dapat dinyatakan sebagai

$$Cv = \lambda v \tag{8}$$

Sekarang akan ditunjukkan bahwa semua solusi  $v$  dengan  $\lambda \neq 0$  terletak pada rentang  $\Phi(x_1), \dots, \Phi(x_m)$ , yaitu

$$\lambda(\Phi(x_k) \cdot v) = (\Phi(x_k) \cdot Cv) \tag{9}$$

dimana;  $k = 1, \dots, m$

$$v = \sum_{i=1}^m a_i \Phi(x_i) \tag{10}$$

substitusi persamaan (7) dan (10) ke dalam persamaan (9), maka

$$m\lambda \sum_{j=1}^m a_j \Phi(x_i) = \sum_{i=1}^m \sum_{j=1}^m a_j \Phi(x_i) K(x_i, x_j) \tag{11}$$

dimana,  $K(x_i, x_j) = \Phi(x_i)^T \Phi(x_j)$  (12)

Berikut [3] merupakan beberapa kernel yang biasa digunakan:

✚ Gauss :  $k(x_i, x_j) = \exp\left(-\frac{1}{2} \left(\frac{\|x_i - x_j\|}{h_0}\right)^2\right)$

✚ Polinomial :  $k(x_i, x_j) = (x_i^T x_j + h_0)^d$

✚ Power :  $k(x_i, x_j) = -\|x_i - x_j\|^d$

### 3. Kernel Polinomial

Menurut [4] salah satu kernel yang banyak digunakan adalah kernel polinomial. Dimana  $h_0$  adalah parameter skala yang akan dipilih. Umumnya,  $h_0$  harus dipilih lebih besar dari jarak khas antar sampel dan lebih kecil dari ukuran daerah penelitian

$$K(x_i, x_j') = (x_i^T x_j' + h_0)^d \tag{13}$$

Sebagai contoh kita menganggap fungsi kernel polinomial dimana  $d = 2$ , dengan dua dimensi  $x = [z_1 \ z_2]^T$  dan  $x' = [z_1'^2 \ z_2'^2]^T$ .

$$\text{Kita memperoleh, } K(x, x') = (x^T x' + h_0)^2 \tag{14}$$

$$\begin{aligned} &= (z_1 z_1' + z_2 z_2' + h_0)^2 \\ &= z_1^2 z_1'^2 + z_2^2 z_2'^2 + h_0^2 + 2z_1 z_1' z_2 z_2' + 2z_1 z_1' h_0 + 2z_2 z_2' h_0 \\ &= \begin{bmatrix} h_0 & \sqrt{2h_0}z_1 & \sqrt{2h_0}z_2 & z_1^2 & z_2^2 & \sqrt{2}z_1 z_2 \end{bmatrix} \begin{bmatrix} h_0 \\ \sqrt{2h_0}z_1' \\ \sqrt{2h_0}z_2' \\ z_1'^2 \\ z_2'^2 \\ \sqrt{2}z_1' z_2' \end{bmatrix} \\ &= \phi(x)^T \phi(x') \end{aligned} \tag{15}$$

Kita dapat melihat bahwa fungsi kernel pemetaan vektor dua dimensi menjadi enam dimensi.

### 4. Metode Analisis

Langkah-langkah dalam metode analisis :

1. Gambar plot yang ada dalam buku “*Multivariate Statistical Methods A PRIMER*” di halaman 105 difoto kembali, kemudian diperbesar dan diprint pada kertas bergaris (*millimeter block*).
2. Buat sumbu koordinat X1 dan X2 untuk menentukan titik koordinat dari X1 dan X2.
3. Titik koordinat pertama yang diperoleh dari sumbu X1 digunakan sebagai data pertama pada peubah X1 dan seterusnya, begitupun dengan X2.
4. Setelah data diperoleh, dilakukan standarisasi dengan mengurangi tiap data dengan nilai rata-rata dari data tersebut dan membaginya dengan simpangan baku.

$$\text{Standarisasi } X_p = \frac{X_{pi} - \mu_p}{s_p} p = 1,2 ; \quad \text{dengan ; } i = 1,2, \dots, n \text{ dan } p = \text{peubah.}$$

5. Digunakan PCA untuk menampilkan data hasil standarisasi dalam *score plot*.
6. Dilakukan *kernel PCA* fungsi polinomial untuk mentransformasi data hasil standarisasi.

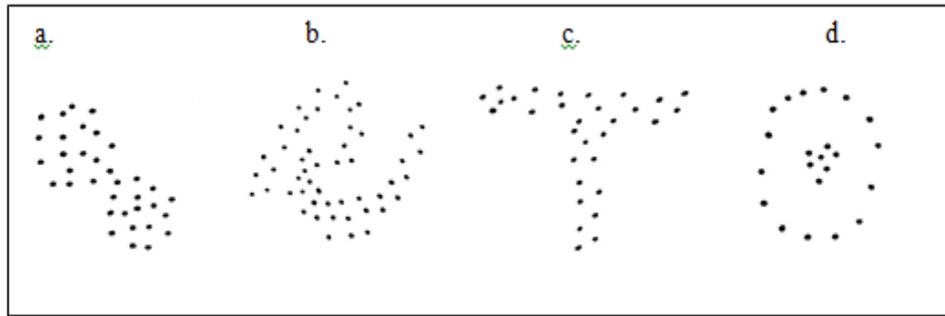
$$K(x_i, x_j') = (x_i^T x_j' + h_0)^d \text{ dengan ; } h_0 = 1 \text{ dan } d = 2$$

7. Setelah itu, dicari skor komponen dari data yang sudah ditransformasi.
8. Ditampilkan matriks plot dari skor komponen data yang distandarisasi dan score component data yang ditransformasi.
9. Interpretasi gambar *matrix plot* yang diperoleh.

### 5. Hasil dan Pembahasan

#### 5.1 Data

Data yang digunakan dalam penelitian ini diambil dari buku “*Multivariate Statistical Methods A PRIMER*” di halaman 105, data berupa plot peubah ganda [5]. Dari buku tersebut dipilih empat gambar plot yang merupakan masalah analisis gerombol. Data-c pada gambar yang ada dibuku digunakan sebagai data-a pada penelitian, demikian juga dengan data-d, data-e, dan data-f yang menjadi data-b, data-c, dan data-d dapat dilihat pada Gambar 4.



Gambar 4. Plot Peubah Ganda Penelitian

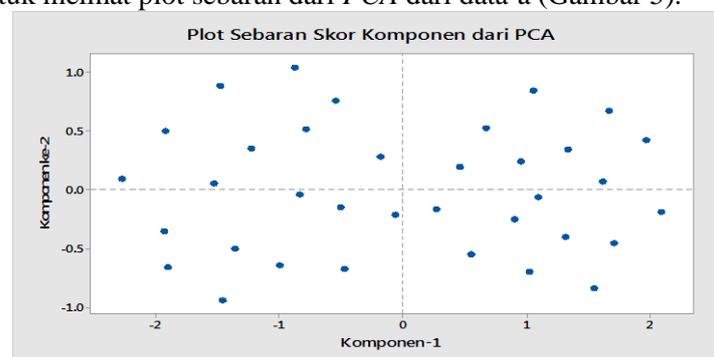
Setiap gambar yang ada dibuku, terlebih dahulu difoto kemudian diperbesar 300 persen dari gambar awal yang ada dibuku, selanjutnya gambar yang ada diprint pada kertas *millimeter block*. Secara sembarang ditarik sumbu koordinat X1 dan X2, tapi dalam penelitian ini untuk keperluan data yang akan digunakan diambil sumbu yang tegak lurus seperti pada bidang koordinat cartesius, sumbu mendatar diberi *label* X1 dan sumbu tegak diberi *label* X2. Untuk memperoleh data dalam penelitian ini dihubungkan sumbu koordinat X1 dan X2. Kemudian dilihat titik-titik koordinat yang merupakan perpotongan sumbu koordinat X1 dan X2 dari keseluruhan individu, hasil perhitungan secara *visual* dengan menghubungkan sumbu koordinat X1 dan X2 pada *millimeter block* diperoleh taksiran untuk angka-angka yang kemudian digunakan sebagai data dalam penelitian ini. Titik koordinat pertama pada sumbu X1 digunakan sebagai data pertama pada peubah X1 begitu seterusnya, juga pada sumbu X2. Posisi titik awal dan besar skala yang digunakan bisa sembarang, tetapi tidak mempengaruhi konfigurasi plot yang ada karena data tersebut akan distandarisasi. Langkah-langkah dilakukan untuk memperoleh data awal pada plot sebaran data-a, juga pada plot sebaran data-b, data-c, dan data-d.

## 5.2 Hasil

Hasil dari *matrix plot* pada sebaran data-a, sebaran data-b, sebaran data-c dan sebaran data-d diperoleh dengan menggunakan bantuan *software* statistika.

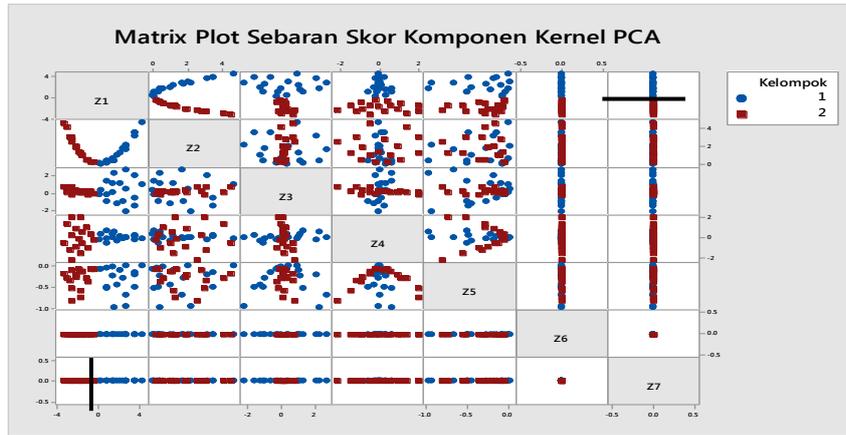
### 5.2.1 Plot Sebaran dari Data-a

Pada plot awal data-a (Gambar 4), terlihat kumpulan titik yang membentuk dua buah lingkaran yang saling berdekatan. Data awal yang telah diperoleh distandarisasi. Selanjutnya dilakukan *PCA* untuk melihat plot sebaran dari *PCA* dari data-a (Gambar 5).

Gambar 5. Plot Sebaran Skor Komponen dari *PCA* untuk Data-a

Plot sebaran skor komponen data hasil standarisasi terlihat masih sama seperti sebaran data awal. Untuk itu dilanjutkan dengan transformasi *kernel PCA* dengan fungsi kernel polinomial. Dan dicari skor komponen dari data hasil standarisasi dan data hasil transformasi. Data skor komponen yang diperoleh, ditampilkan dalam *matrix plot*.

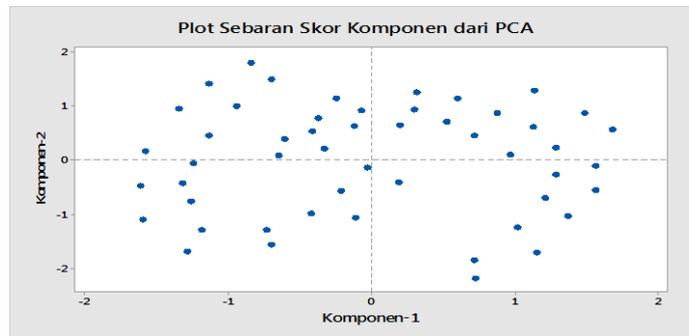
Untuk plot sebaran data-b, data-c dan data-d menggunakan langkah-langkah yang sama seperti pada plot sebaran data-a.



Gambar 6. Matrix Plot Sebaran Skor Komponen *Kernel PCA* untuk Data-a

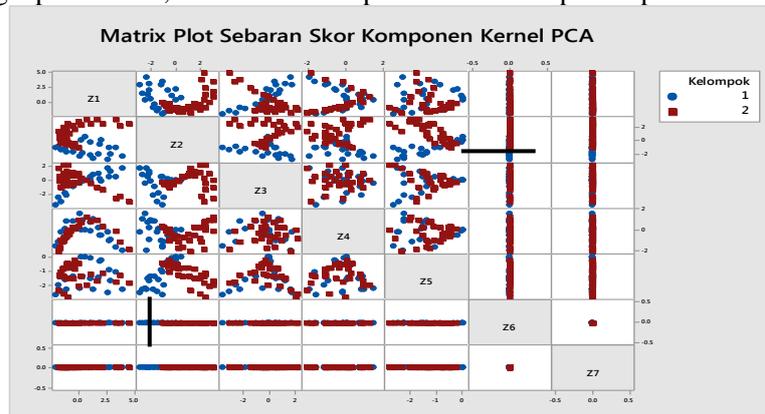
### 5.2.2 Plot Sebaran dari Data-b

Data plot awal sebaran data-b yang diperoleh selanjutnya distandarisasi. Data hasil standarisasi, ditampilkan pada plot sebaran skor komponen.



Gambar 7. Plot Sebaran Skor Komponen dari *PCA* untuk Data-b

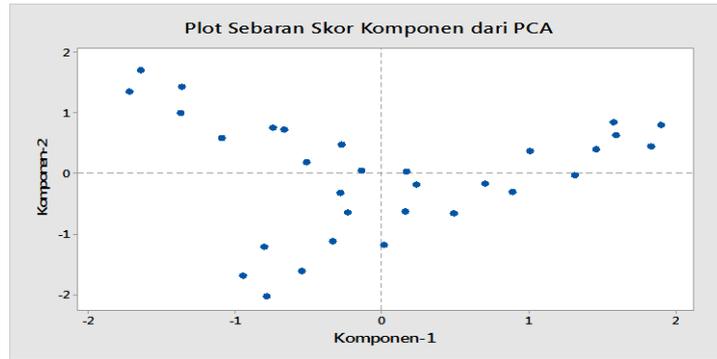
Pada sebaran skor komponen dari *PCA* untuk data-b (Gambar 7), belum dapat dikelompokkan dengan jelas sehingga dilakukan transformasi dari data hasil standarisasi. Data hasil transformasi *Kernel PCA* fungsi polinomial, dicari skor komponen dan ditampilkan pada matriks plot



Gambar 8. Matrix Plot Sebaran Skor Komponen *Kernel PCA* untuk Data-b

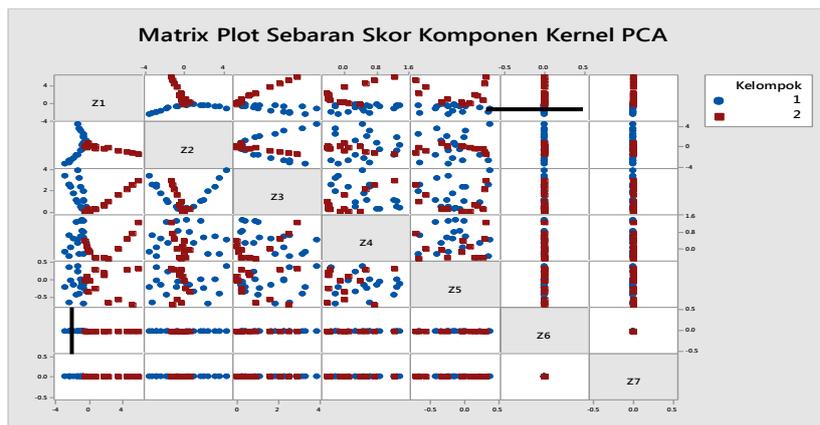
### 5.2.3 Plot Sebaran dari Data-c

Plot awal sebaran data-c (Gambar 4) terlihat seperti sebaran titik yang membentuk huruf T. Data awal yang diperoleh, kemudian distandarisasi. Data hasil standarisasi, ditampilkan pada plot sebaran skor komponen dari *PCA*.



Gambar 9. Plot Sebaran Skor Komponen dari *PCA* untuk data-c

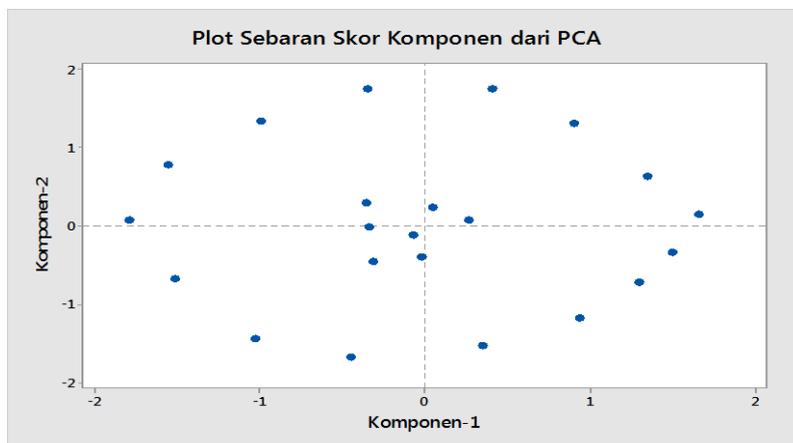
Dari Gambar 9 terlihat sebaran titik yang ada membentuk huruf T yang terbalik dan seperti menyamping dan masih sama seperti plot awal. Dilakukan transformasi dan dicari skor komponen yang kemudian ditampilkan pada matrix plot untuk data-c (Gambar 10).



Gambar 10. Matrix Plot Sebaran Skor Komponen *Kernel PCA* untuk Data-c

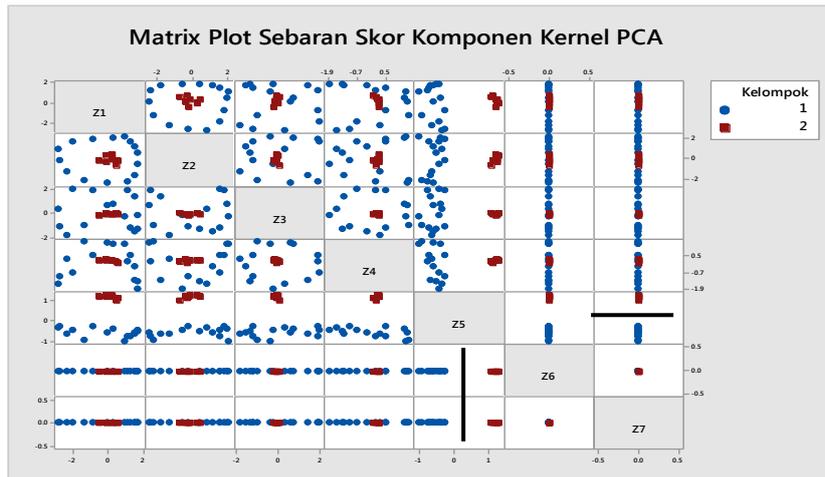
### 5.2.4 Plot Sebaran dari Data-d

Dari gambar sebaran plot awal data-d (Gambar 4) terlihat sebaran titik yang terbagi menjadi dua kelompok, kelompok pertama berupa lingkaran dan kelompok kedua berupa kumpulan titik yang terletak didalam lingkaran. Data distandarisasi, kemudian ditampilkan pada *score plot* dengan menggunakan PCA.



Gambar 11. Plot Sebaran Skor Komponen dari *PCA* untuk data-d

Dari plot diatas (Gambar 11), terlihat bahwa sebaran data yang masih sama seperti sebaran data awal hanya saja pada kedua sisi terlihat runcing dan belum dapat dipisahkan secara garis linier. Dilakukan transformasi, lalu dicari skor komponen dari data hasil standarisasi dan hasil transformasi dan ditampilkan pada matrix plot.



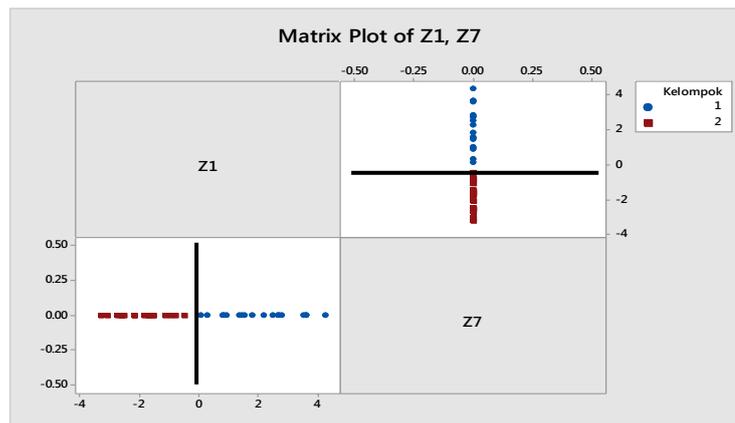
Gambar 12. Matrix Plot Sebaran Skor Komponen *Kernel PCA* untuk Data-d

### 5.3 Pembahasan

Jika terdapat satu saja komponen pada matrix plot yang kelompoknya dapat dipisahkan secara garis linier, berarti masalah pengelompokan tersebut dapat diselesaikan dengan fungsi kernel polinomial.

#### 5.3.1 Matrix Plot Skor Komponen *Kernel PCA* untuk Data-a

Dari matriks plot sebaran individual dapat dibuat suatu garis linier yang dapat mendiskriminasi sebaran individu kedalam dua kelompok yang berbeda secara jelas. Terlihat pada matrix plot sebaran skor komponen *kernel PCA* untuk data-a (Gambar 6) pada setiap komponen utama pertama dapat dipisahkan dua kelompok secara garis linier (Gambar 13).

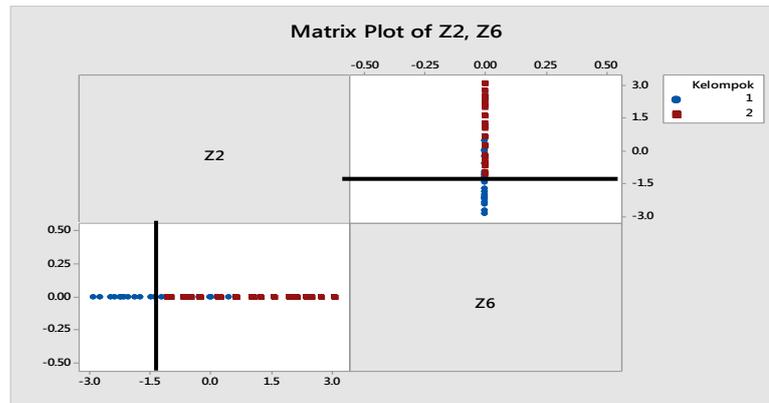


Gambar 13. Matrix Plot Komponen Z1 dan Komponen Z7

Pada komponen pertama (Z1) dan komponen ketujuh (Z7) terlihat individu yang ada telah mengelompok seperti sebuah garis lurus dan secara jelas terbagi menjadi dua bagian, yaitu kelompok yang berwarna biru (kelompok 1) berada dan kelompok merah (kelompok 2) berada dibagian bawah, sehingga dapat ditarik garis pemisah yang memisahkan kelompok 1 dan kelompok 2 secara garis linier.

#### 5.3.2 Matrix Plot Skor Komponen *Kernel PCA* untuk Data-b

Pada matrix plot sebaran skor komponen *kernel PCA* untuk data-b (Gambar 8), garis linier dapat digunakan untuk mendiskriminasi sebaran individu kedalam dua kelompok yang berbeda, walaupun pada sejumlah kelompok kecil perlu ketelitian khusus untuk memprediksi hasil diskriminasi tersebut, seperti pada Gambar 14.

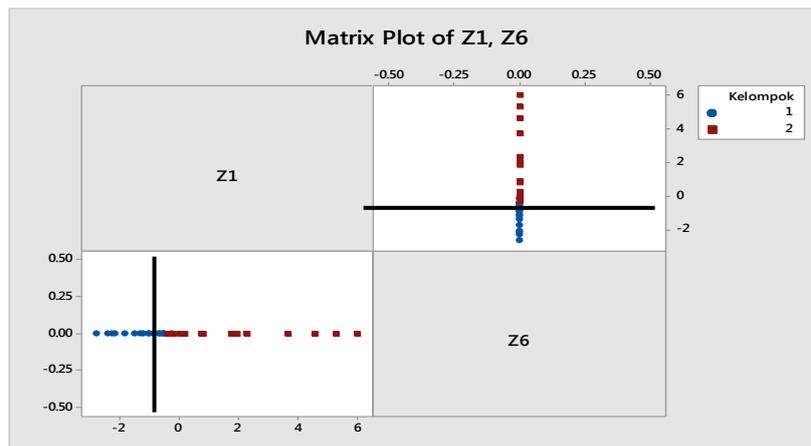


Gambar 14. Matrix Plot Komponen Z2 dan Komponen Z6

Jika kita menarik garis pemisah antara komponen Z2 dan komponen Z6 (Gambar 14), maka akan terdapat beberapa individu pada kelompok merah (kelompok 2) yang menyebar pada kelompok biru (kelompok 1) begitu juga pada kelompok merah (kelompok 2). Hal ini dikarenakan jarak individu dalam dua kelompok sangat dekat sehingga terdapat individu yang bertumpang tindih.

### 5.3.3 Matrix Plot Skor Komponen Kernel PCA untuk Data-c

Dari matrix plot sebaran skor komponen *kernel PCA* untuk data-c (Gambar 10) dapat dibuat suatu garis linier yang dapat mendiskriminasi sebaran-sebaran individu kedalam dua kelompok yang berbeda, bisa dilihat pada Gambar 15. Dengan menarik garis pemisah pada komponen pertama (Z1) dan komponen keenam (Z6), maka dapat dibuat pengelompokan secara linier antara kelompok yang berwarna biru (kelompok 1) dan kelompok yang berwarna merah (kelompok 2) dengan baik.

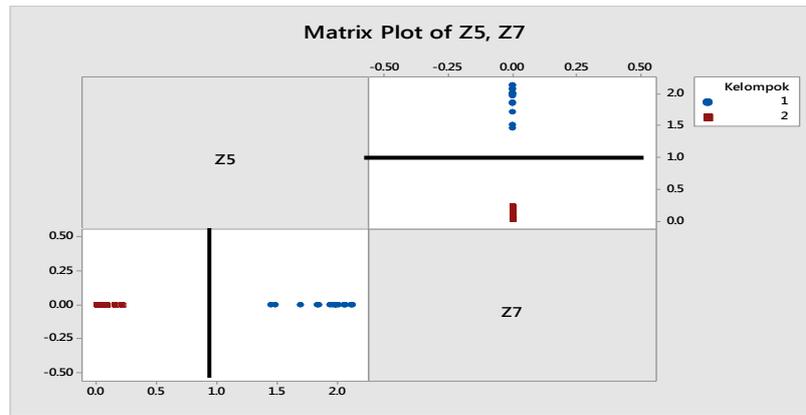


Gambar 15. Matrix Plot Komponen Z1 dan Komponen Z6

### 5.3.4 Matrix Plot Skor Komponen Kernel PCA untuk Data-d

Dari matrix plot sebaran skor komponen *kernel PCA* untuk data-d pada Gambar 12, terlihat dua kelompok yang dapat dipisahkan dengan jelas dengan menarik suatu garis linier. Suatu garis linier dapat mendiskriminasi sebaran individu kedalam dua kelompok yang berbeda. Misalnya, dapat dilihat pada Gambar 16.

Pada komponen utama kelima (Z5) dan komponen utama ketujuh (Z7) jika dibuat garis pemisah secara linier, maka dapat mengelompokkan kelompok 1 (warna biru) dengan kelompok 2 (warna merah) dengan sangat jelas. Hal ini dikarenakan skala antara kelompok 1 dan kelompok 2 pada plot awal sebaran data-d yang agak jauh.



Gambar 16. Matrix Plot Komponen Z5 dan Komponen Z7

## 6. Kesimpulan

Penggunaan *kernel PCA (KPCA)* fungsi polinomial sangat membantu dalam menyelesaikan masalah pengelompokan plot peubah ganda yang belum dapat dikelompokkan atau dipisahkan secara garis linier. Dengan menggunakan *kernel PCA* fungsi polinomial, pada matrix plot sebaran skor komponen *Kernel PCA* untuk sebaran data-a, suatu garis linier dapat mendiskriminasi sebaran individu kedalam dua kelompok yang berbeda begitu juga pada sebaran data-b, data-c dan data-d.

## 7. Daftar Pustaka

- [1] Djakaria I., S. Guritno, dan S.H. Kartiko. 2010. Visualisasi Data Iris Menggunakan Analisis Komponen Utama dan Analisis Komponen Utama Kernel. *Jurnal ILMU DASAR*. 11(1):31-38.
- [2] Schölkopf B., S. Alexander & R.M. Klaus. 1996. *Nonlinear Component Analysis as a Kernel Eigenvalue Problem*. Technical Report No. 44, Nanyang Technological University.
- [3] Thamara, P. 2014. Analisis Komponen Utama Kernel: Suatu Studi Eksplorasi Pembakuan Peubah [Skripsi]. Departemen Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam IPB, Bogor.
- [4] Nielsen, A.A and M.J. Canty. 2008. *Kernel Principal Component Analysis for Change Detection*. Proceeding of SPIE, Vol.7109: 1-10.
- [5] Manly, B.F.J. 1986. *Multivariate Statistical Methods A PRIMER*. Chapman and Hall, New York.