

PERHITUNGAN NILAI EIGEN

Nilai eigen merupakan salah satu konsep penting dalam aljabar linear. Konsep nilai eigen sering dianggap sulit karena melibatkan matriks dan operasi matematis. Padahal secara sederhana, nilai eigen menunjukkan seberapa besar suatu transformasi memperbesar atau memperkecil suatu vektor tertentu.

Dalam bidang data science, nilai eigen dan vektor eigen banyak digunakan pada metode Principal Component Analysis (PCA). PCA digunakan untuk menyederhanakan data yang memiliki sangat banyak variabel menjadi beberapa komponen utama tanpa menghilangkan informasi penting. Misalnya pada perusahaan e-commerce yang memiliki data pelanggan berupa umur, pendapatan, frekuensi belanja, jumlah transaksi, dan lama penggunaan aplikasi. Jika seluruh variabel dianalisis secara langsung, proses komputasi menjadi lebih kompleks dan sulit divisualisasikan. Oleh karena itu digunakan PCA untuk mengurangi dimensi data.

Pada proses PCA, nilai eigen digunakan untuk menentukan seberapa besar informasi yang disimpan oleh setiap komponen utama, sedangkan vektor eigen digunakan untuk menentukan arah pembentukan komponen tersebut. Komponen dengan nilai eigen terbesar dianggap paling penting karena mampu menjelaskan variasi data paling besar. Hasil analisis ini dapat membantu perusahaan memahami pola pelanggan, melakukan segmentasi pasar, hingga meningkatkan sistem rekomendasi produk.

KONSEP DASAR NILAI EIGEN

1. Pengertian Nilai Eigen

Jika terdapat matriks A dan vektor x :

$$Ax = \lambda x$$

Maka:

- A = matriks
- x = vektor eigen

- λ = nilai eigen

Artinya:

Ketika matriks A dikalikan dengan vektor x, arah vektor tetap sama namun panjangnya berubah sebesar nilai eigen.

Rumus Nilai Eigen

Nilai eigen diperoleh dari:

$$|A - \lambda I| = 0$$

Keterangan:

- A = matriks
- I = matriks identitas
- λ = nilai eigen

STUDI KASUS

ANALISIS NILAI EIGEN UNTUK MENENTUKAN FAKTOR DOMINAN PENJUALAN di SUPERMARKET

1. Contoh Kasus

Analisa sebuah perusahaan supermarket besar yang memiliki banyak cabang di berbagai kota. Manajemen ingin memahami faktor utama yang memengaruhi performa penjualan.

Langkah awal yang dilakukan Perusahaan dengan mengumpulkan data dari 25 cabang supermarket dengan variabel:

1. Jumlah pengunjung
2. Pengeluaran promosi
3. Jumlah produk tersedia
4. Pendapatan harian

Pihak manajemen mengalami kesulitan karena data memiliki banyak variabel yang saling berkaitan.

Mereka ingin mengetahui:

- Faktor mana yang paling dominan?
- Variabel mana yang paling berpengaruh?
- Bagaimana menyederhanakan data menjadi informasi utama?

Untuk menyelesaikan masalah tersebut digunakan konsep nilai eigen pada PCA.

2. Dataset

Cabang	Pengunjung	Promosi	Produk	Pendapatan
1	230	15	420	58
2	250	16	430	61
3	245	18	440	62
4	260	20	450	65
5	280	22	470	70
6	290	24	480	72
7	300	25	490	75
8	310	27	500	77
9	320	28	510	80
10	330	29	520	82
11	340	30	530	84
12	350	31	540	86
13	360	32	550	88
14	370	34	560	90
15	380	35	570	92
16	390	36	580	94
17	400	38	590	96
18	410	39	600	98
19	420	40	610	100
20	430	42	620	102

21	440	43	630	104
22	450	45	640	106
23	460	46	650	108
24	470	47	660	110
25	480	48	670	112

LANGKAH PENYELESAIAN DENGAN R

1. Membuat Dataset

```
# Membuat data supermarket
supermarket <- data.frame(
  Pengunjung = c(230,250,245,260,280,
                 290,300,310,320,330,
                 340,350,360,370,380,
                 390,400,410,420,430,
                 440,450,460,470,480),
  Promosi = c(15,16,18,20,22,
              24,25,27,28,29,
              30,31,32,34,35,
              36,38,39,40,42,
              43,45,46,47,48),
  Produk = c(420,430,440,450,470,
             480,490,500,510,520,
             530,540,550,560,570,
             580,590,600,610,620,
             630,640,650,660,670),
  Pendapatan = c(58,61,62,65,70,
                 72,75,77,80,82,
                 84,86,88,90,92,
                 94,96,98,100,102,
                 104,106,108,110,112)
)

# Menampilkan data
supermarket
```

Data disusun dalam bentuk data frame. Setiap kolom merepresentasikan variabel berbeda.

2. Standardisasi Data

```
# Standardisasi data
data_scaled <- scale(supermarket)

# Menampilkan hasil
head(data_scaled)
```

Pada PCA, data harus distandardisasi karena setiap variabel memiliki skala berbeda.

Contoh:

- Pengunjung : ratusan
- Promosi: puluhan
- Pendapatan: ratusan

Jika tidak distandardisasi, variabel dengan angka besar akan mendominasi analisis.

Fungsi `scale()` digunakan untuk:

- Mengurangi mean menjadi 0
 - Mengubah standar deviasi menjadi 1
-

3. Membuat Matriks Korelasi

```
# Matriks korelasi
correlation_matrix <- cor(data_scaled)

# Menampilkan matriks korelasi
correlation_matrix
```

Matriks korelasi menunjukkan hubungan antar variabel. Nilai mendekati 1 berarti hubungan sangat kuat.

4. Menghitung Nilai Eigen

```
# Menghitung eigen
eigen_result <- eigen(correlation_matrix)

# Menampilkan hasil eigen
eigen_result
```

Fungsi `eigen()` digunakan untuk menghitung:

1. Nilai eigen
2. Vektor eigen

Nilai eigen menunjukkan besarnya informasi yang disimpan oleh masing-masing komponen.

5. Menampilkan Nilai Eigen

```
# Menampilkan nilai eigen
eigen_result$values
```

Semakin besar nilai eigen, semakin penting komponen tersebut. Biasanya komponen dengan nilai eigen > 1 dianggap penting.

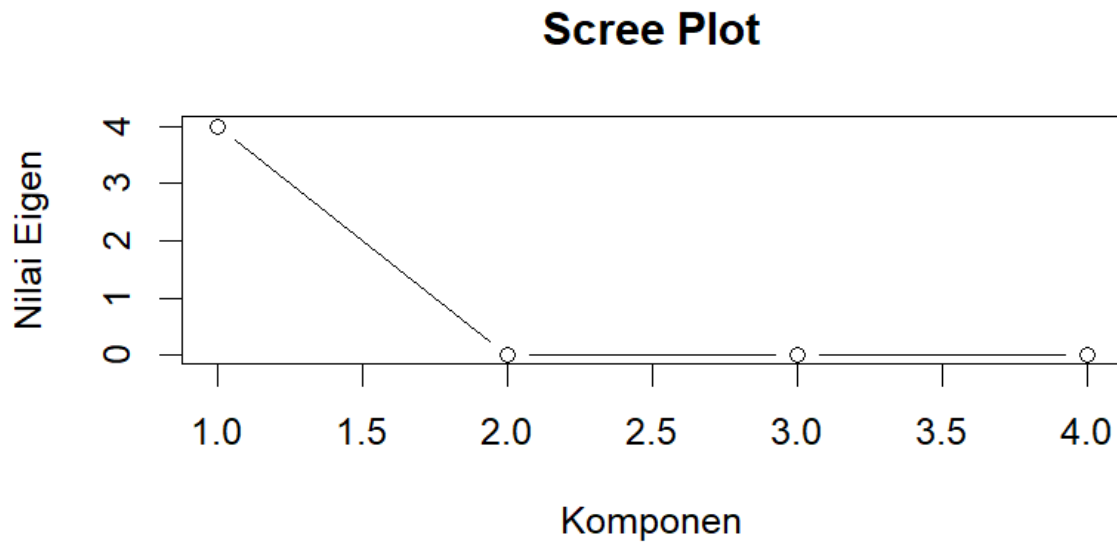
6. Menampilkan Vektor Eigen

```
# Menampilkan vektor eigen
eigen_result$vectors
```

Vektor eigen menunjukkan kontribusi masing-masing variabel terhadap komponen utama.

7. Membuat Scree Plot

```
# Membuat scree plot
plot(eigen_result$values,
     type = "b",
     main = "Scree Plot",
     xlab = "Komponen",
     ylab = "Nilai Eigen")
```



Scree plot digunakan untuk melihat jumlah komponen penting.

Jika grafik turun tajam setelah komponen tertentu, maka komponen sebelumnya dianggap paling dominan.

8. Menggunakan PCA

```
# PCA
pca_result <- prcomp(supermarket,
                     scale = TRUE)

# Ringkasan PCA
summary(pca_result)
```

Fungsi `prcomp()` digunakan untuk melakukan PCA.

Parameter `scale = TRUE` digunakan agar data otomatis distandardisasi.

ANALISIS HASIL

Berdasarkan hasil nilai eigen:

1. Komponen pertama memiliki nilai eigen terbesar.
2. Artinya sebagian besar variasi data dijelaskan oleh komponen pertama.
3. Variabel pengunjung, promosi, dan produk kemungkinan memiliki hubungan kuat dengan pendapatan.

Manajemen dapat menggunakan hasil ini untuk:

- Fokus pada strategi promosi
- Meningkatkan jumlah produk
- Mengoptimalkan pengalaman pelanggan

Daripada menganalisis semua variabel secara terpisah, PCA menyederhanakan data menjadi beberapa faktor utama.

INTERPRETASI BISNIS

Hasil PCA membantu perusahaan memahami bahwa:

- Cabang dengan pengunjung tinggi cenderung memiliki pendapatan tinggi.
- Promosi yang lebih besar meningkatkan transaksi.
- Jumlah produk yang lengkap meningkatkan minat pembeli.

Dengan demikian perusahaan dapat mengalokasikan anggaran lebih efektif.

KESIMPULAN

Nilai eigen merupakan konsep penting dalam aljabar linear dan data science.

Melalui nilai eigen, data yang kompleks dapat disederhanakan menjadi informasi utama.

Dalam dunia nyata, konsep ini sangat berguna untuk:

- Analisis data besar
- Reduksi dimensi
- Machine learning
- Pengambilan keputusan bisnis

Penggunaan R mempermudah proses perhitungan matriks, nilai eigen, visualisasi, dan PCA secara otomatis.