Sivas Cumhuriyet Üniversitesi
Bilgisayar Mühendisliği Bölümü

Bilgisayar Destekli Lineer Cebir

Dr. Öğr. Üyesi Fırat İsmailoğlu

Özdeğer - Özvektör

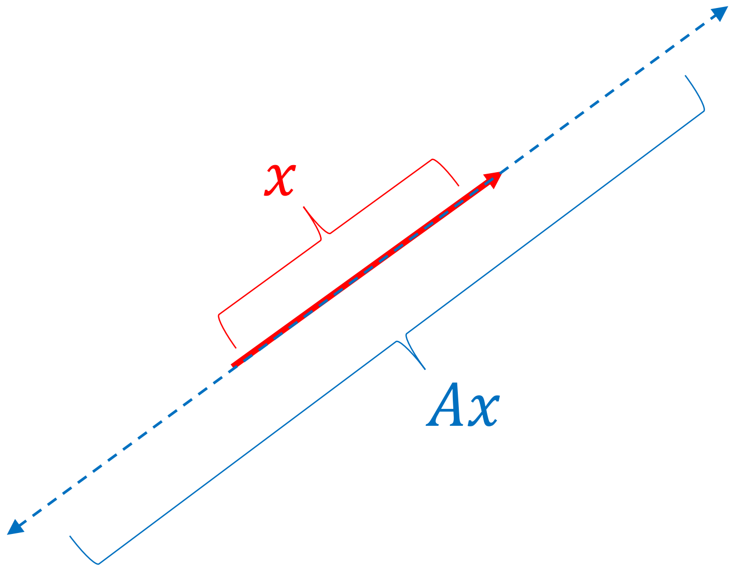
Özdeğer – Özvektör (Eigenvalue – Eigenvector)

Şimdiye kadar hep $Ax = b$ şeklinde lineer denklem sistemlerinin çözümü için uğraştık. Şimdi ise denklemin sağ tarafında b vektörü yerine, A matrisi ile çarpılan x vektörü ile aynı doğrultuda yer alan λx vektörü yer alsın. Bu, bazı özel sonuçlar doğurur.

$$\boxed{Ax = \lambda x}$$

eşitliğini sağlayan λ değerlerine A'nın öz değerleri (eigen values), bu λ'lara karşılık gelen x vektörlerine ise A'nın öz vektörleri (eigen vectors) diyeceğiz.

Öncelikle λ ve x'lerin nasıl bulunacağına bakalım, daha sonra bunların neyi ifade ettiklerini, nerelerde kullanıldığını inceleyelim.



Özdeğerlerin Bulunması

$$Ax = \lambda x$$
$$Ax - \lambda x = \mathbf{0}$$
$$(A - \lambda I)x = \mathbf{0}$$

Burada I birim matristir. En son elde edilen $(A - \lambda I)x = \mathbf{0}$ eşitliğinde eğer x'in katsayısı olan $(A - \lambda I)$ matrisinin tersi varsa $x = (A - \lambda I)^{-1}\ \mathbf{0}$ olur, bu ise x'in 0 vektörü $(\mathbf{0})$ yani tüm elemanları 0 olan vektör olması demektir. Biz böyle bir vektör aramıyoruz. Bunun olmaması için $(A - \lambda I)^{-1}$ olmaması gerekir, bu ise ancak bu matrisin determinantinin 0 olması ile mümkündür. O halde bu matrisin dterminantini 0 yapan λ değerleri arıyoruz:

A'nın öz vektörleri:

$$\det(A - \lambda I) = 0$$

yapan λ değerleridir. Şimdi bir örnek üzerinden bunları nasıl bulacağımıza bakalım.

ör. $A = \begin{bmatrix} 2 & -4 \\ -1 & -1 \end{bmatrix}$ matrisinin öz değerleri nelerdir?

çözüm. Önce $(A - \lambda I)$ matrisini oluştururalım:

$$A - \lambda I = \begin{bmatrix} 2 & -4 \\ -1 & -1 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} 2 & -4 \\ -1 & -1 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$$
$$= \begin{bmatrix} 2 - \lambda & -4 \\ -1 & -1 - \lambda \end{bmatrix}$$

$\begin{bmatrix} 2-\lambda & -4 \\ -1 & -1-\lambda \end{bmatrix}$ matrisinin determınantını 0 'a eşitleyen λ değerleri aradığımız özdeğerlerdir.

$$det\left(\begin{bmatrix} 2-\lambda & -4 \\ -1 & -1-\lambda \end{bmatrix}\right) = 0$$
$$(2-\lambda)(-1-\lambda) - 4 = 0$$
$$\lambda^2 - \lambda - 6 = 0$$

Bu denklemin kökleri $\lambda_1 = 3$ ve $\lambda_2 = -2$ 'dir.

Özvektörlerin Bulunması

Bunun için bulduğumuz λ değerlerini $(A - \lambda I)x = \mathbf{0}$ eşitliğinde yerine koyup bunlara karşılık gelen x vektörlerini bulacağız.

Bir önceki örnekte gördüğümüz $A = \begin{bmatrix} 2 & -4 \\ -1 & -1 \end{bmatrix}$ matrisi için özvektörleri bulalım.

$\lambda_1 = 3$ için:

$$\left(\begin{bmatrix} 2 & -4 \\ -1 & -1 \end{bmatrix} - \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}\right)\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
$$\begin{bmatrix} -1 & -4 \\ -1 & -4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
$$-x_1 - 4x_2 = 0$$
$$-x_1 - 4x_2 = 0$$

Yani sonuç olarak bir (bağımsız) denklemimiz iki tane bilinmeyenimiz var. Bu da demektir ki x_1 ve x_2 bir değişken cinsinden bulunacak.

$x_2 = t$ dersek, $x_1 = -4t$ olur. Şu halde $\lambda_1 = 3$ e karşılık gelen öz vektörler $t\begin{bmatrix}-4\\1\end{bmatrix}$ tipindedir.

Şimdi ikinci grup öz özvektörleri bulalım. $\lambda_2 = -2$ için

$$\left(\begin{bmatrix}2 & -4\\-1 & -1\end{bmatrix} - \begin{bmatrix}-2 & 0\\0 & -2\end{bmatrix}\right)\begin{bmatrix}x_1\\x_2\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix}$$

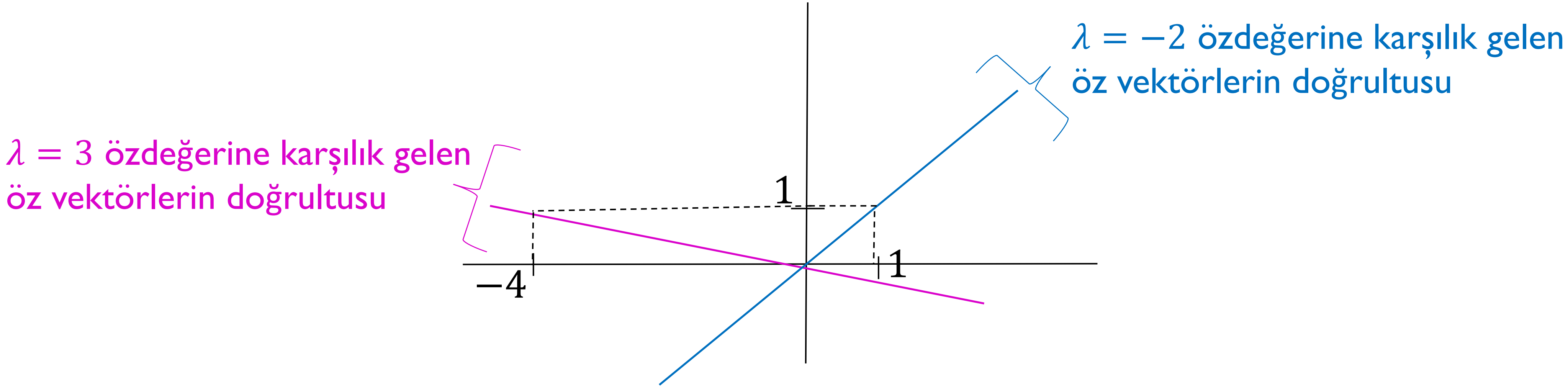
$$\begin{bmatrix}4 & -4\\-1 & 1\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix}$$

$$4x_1 - 4x_2 = 0$$

$$-x_1 + x_2 = 0$$

Buradan x_1=x_2 bulunur. O halde $\lambda_2 = -2$ ye karşılık gelen öz vektörler $t\begin{bmatrix}1\\1\end{bmatrix}$ tipindedir

Bulduğumuz öz vektörlerin birinci grubundakiler aşağıda gösterilen pembe doğrultuda; ikinci gruptakiler ise mavi doğrultuda gösterilmiştir.



ör. $A = \begin{bmatrix} 5 & 8 & 16 \\ 4 & 1 & 8 \\ -4 & -4 & -11 \end{bmatrix}$ matrisinin öz değer ve öz vektörlerini bulunuz.

$$A - \lambda I = \begin{bmatrix} 5-\lambda & 8 & 16 \\ 4 & 1-\lambda & 8 \\ -4 & -4 & -11-\lambda \end{bmatrix}$$

olur. Bu matrisin determinantını 0'a eşitleyelim:

$$\det(A - \lambda I) = (\lambda - 1)(\lambda + 3)^2 = 0$$

olduğundan aradığımız öz değerler $\lambda_1 = 1, \lambda_2 = -3$ ve $\lambda_3 = -3$ olarak bulunur.

$\lambda_1 = 1$ 'e karşılık gelen öz vektörleri bulalım.

$$\left(\begin{bmatrix} 5 & 8 & 16 \\ 4 & 1 & 8 \\ -4 & -4 & -11 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right)\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\left(\begin{bmatrix} 4 & 8 & 16 \\ 4 & 0 & 8 \\ -4 & -4 & -12 \end{bmatrix}\right)\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$4x_1 + 8x_2 + 16x_3 = 0$$

$$4x_1 + 8x_3 = 0$$

$$-4x_1 - 4x_2 - 12x_3 = 0$$

Elementer satır işlemleri ile aradığımız bilinmeyenleri bulalım:

$$\left[\begin{array}{ccc|c} 4 & 8 & 16 & 0 \\ 4 & 0 & 8 & 0 \\ -4 & -4 & -12 & 0 \end{array}\right] \overset{R_1 \leftarrow (1/4)R_1}{\equiv} \left[\begin{array}{ccc|c} 1 & 2 & 4 & 0 \\ 4 & 0 & 8 & 0 \\ -4 & -4 & -12 & 0 \end{array}\right] \overset{\substack{R_2 \leftarrow -4R_1 + R_2 \\ R_3 \leftarrow 4R_1 + R_3}}{\equiv} \left[\begin{array}{ccc|c} 1 & 2 & 4 & 0 \\ 0 & -8 & -8 & 0 \\ 0 & 4 & 4 & 0 \end{array}\right]$$

$$\overset{R_3 \leftarrow (1/2)R_2 + R_3}{\equiv} \left[\begin{array}{ccc|c} 1 & 2 & 4 & 0 \\ 0 & -8 & -8 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \overset{R_2 \leftarrow (1/8)R_2}{\equiv} \left[\begin{array}{ccc|c} 1 & 2 & 4 & 0 \\ 0 & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$-x_2 - x_3 = 0$ bulunur. $x_3 = t$ dersek $x_2 = -t$ olur. Ilk denklemde yerine yazılırsa $x_1 = -2t$ olur.

Şu halde $\lambda_1 = 1$ 'e karşılık gelen özvektörler $t\begin{bmatrix}-2\\-1\\1\end{bmatrix}$ $(t \in \mathbb{R})$ tipindedir. Örnek olarak t'ye 2 alarak verilen

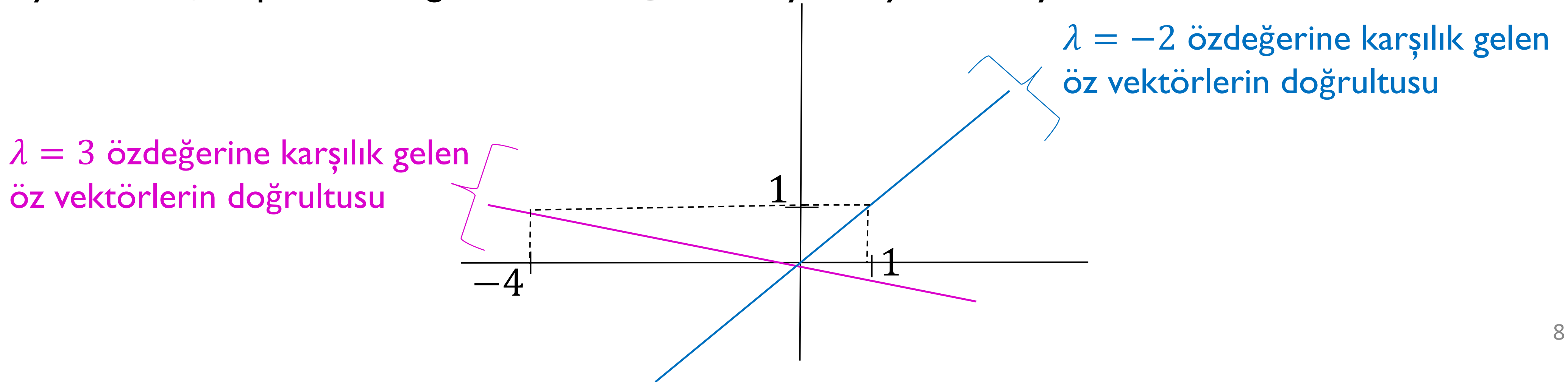
$A = \begin{bmatrix}5 & 8 & 16\\4 & 1 & 8\\-4 & -4 & -11\end{bmatrix}$ matrisinin bir özvektörü $\begin{bmatrix}-4\\-2\\2\end{bmatrix}$ olur. Bu matrisle bu vektörü çarparsak

$$\begin{bmatrix}5 & 8 & 16\\4 & 1 & 8\\-4 & -4 & -11\end{bmatrix}\begin{bmatrix}-4\\-2\\2\end{bmatrix} = 1 \cdot \begin{bmatrix}-4\\-2\\2\end{bmatrix}$$

olur.

Özdeğer – Özvektör Kavramlarının Geometrik Yorumu

Özvektörler matrsilerin uzayda uzandıkları/yayıldıkları doğrultuları gösterirler. Bunlara karşılık gelen özdeğerler ise o doğrultunun ne kadar önemli olduğunu verir. Örneğin daha önce gördüğümüz $A = \begin{bmatrix}2 & -4\\-1 & -1\end{bmatrix}$ matrisi için, bu matris uzayda daha çok pembe ile gösterilen doğrultuda yer alıyordur diyebiliriz.



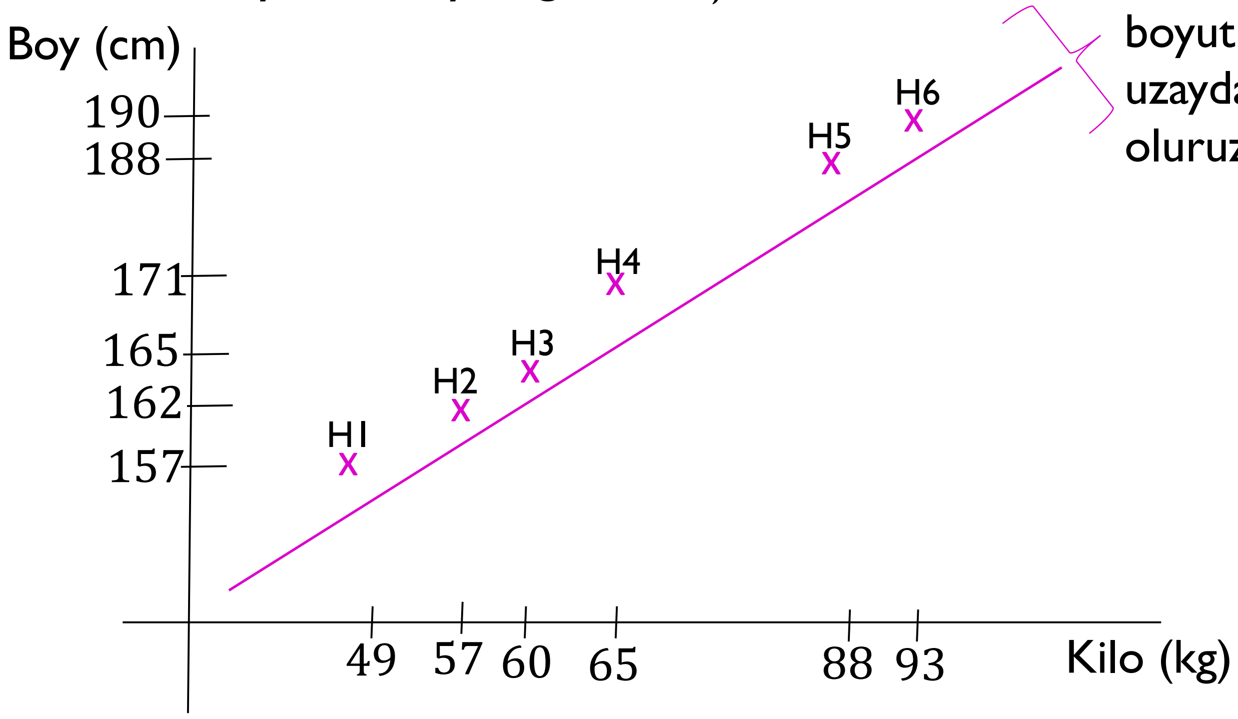
Temel Bileşen Analizi (Principal Component Analysis - PCA)

Özdeğerler-Özvektörler biligisayar-veri bilimlerinde çok yoğun olarak kullanılır. Bunlardan en çok kullanılan uygulması PCA'dır. PCA kısaca veri matrsinin boyunun azaltılmasında (kolon sayının düşürülmesinde) kullanlır. Boyutun fazla olduğu veride (kolon sayısının fazla olduğu yerde) veriler (satırlar) birbirinden çok uzaktatdır, bu durumda saağlıklı bir sekilde yakınlık-uzaklık hesabı yapamayız; ayrıca veriyi görselleştiremeyiz. Literatürde bu duruma boyutun laneti (curse of dimensionality) denir. İşte bu yüzden herhangi bir boyut azaltma (dimensionality reduction) metodu ile verinin boyutunu azaltmamız gerekir. Bütün boyut azaltma metotlarinda asıl amaç orijinal uzaydaki (boyutu azaltilmamış olan) mesafeleri korumaya çalışmaktır.

Basit olması açısından diyelimki aşağıdaki gibi iki boyutlu (iki kolonlu) bir veri matrisimiz olsun. Biz bu veriyi bir boyuta (bir doğru üzerine) indirelim.

Hasta No	Kilo (Kg)	Boy (Cm)
H1	49	157
H2	57	162
H3	60	165
H4	65	171
H5	88	188
H6	93	190

Bu hastaları iki boyutlu uzayda görselleşirisek:



Bu noktaları (hastaları) böyle bir doğru üzerinde tek boyutlu temsil edebilirsek hastaların orijinal iki boyutlu uzaydaki birbirine olan uzaklıklarını mümkün olduğunca korumuş oluruz. Şimdi bu doğrunun nasıl bulunacağına bakalım.

Aradağımız doğru elimizdeki veri matrisinin birinci temel bileşeni (principal component) dir. Bu doğru elimizdeki matrise karşılık gelen *kovaryans* matrsinin birinci özvektörüdür.

Kovaryans matrisinin diagonal elemanlari kolonlarin varyanslarini-yani kolonun merkezlerinden ne kadar dağildiğinin sayisal degerini, diagonal olmayan elemanlar ise kolonlarin kovaryanslari verir, yani kolonlarin birbirine göre ilişkilerinin – beraber artip/azalmadiklarinin- sayisal degerini verir.

Önce varyansı hesaplayalım. Varyans bir sayı topluluğunun merkezlerinden (ortalamalarından) ortalama ne kadar saptığının sayısal değeridir. σ^2 ile gösterilir. Basitçe her sayının merkezden farkı alinarak bu farklarin karesi alinir ki varyans negatif cikmasin daha sonra bu farklar toplanarak ortalamasi alinir.

Not standart sapma, varyansın kareköküdür: standart sapma=$\sqrt{\sigma^2} = \sigma$

Örnekteki Kilo kolonu için varyans:

$$\sigma_{kilo}^2 = \frac{1}{6}\left((49-68)^2 + (57-68)^2 + (60-68)^2 + (65-68)^2 + (88-68)^2 + (93-68)^2\right) = 315$$

Boy kolonu için varyans:

$$\sigma_{boy}^2 = \frac{1}{6}\left((157-172)^2 + (162-172)^2 + (165-172)^2 + (171-172)^2 + (188-172)^2 + (190-172)^2\right) = 190$$

Burada 68 kilo kolonunun ortalaması, 172 ise boy kolonunun ortalamasıdır.

Kilo ve Boy'un kovaryansini hesaplayalım.

$$\sigma_{kilo-boy}^2 = \frac{1}{6}\left((49-68)(157-172) + (57-68)(162-172) + (60-68)(165-172)\right) + (65-68)(171-172) + (88-68)(188-172) + (93-68)(190-172)) = 244$$

$\sigma_{boy-kilo}^2$'yu ayrica hesaplamaya ihtiyacimiz yoktur; çarpimin degisme özelliginden bu $\sigma_{kilo-boy}^2$'a eşittir.

Kovaryans hesabında dikkat etmemiz gereken şudur. Kovaryansın yani benzerliğin artması için kovaryansi oluşturan toplamlardan pozitif degerler gelmesi gerekir. Bu ise ancak çarpımların pozitif bir sonuç dogurmasiyla iliskilidir. Bilindigi gibi bir carpimin pozitif olmasi icin carpanlarin ayni işaretli olmasi gerekir; ikisi de pozitif yada ikisi de negatif.

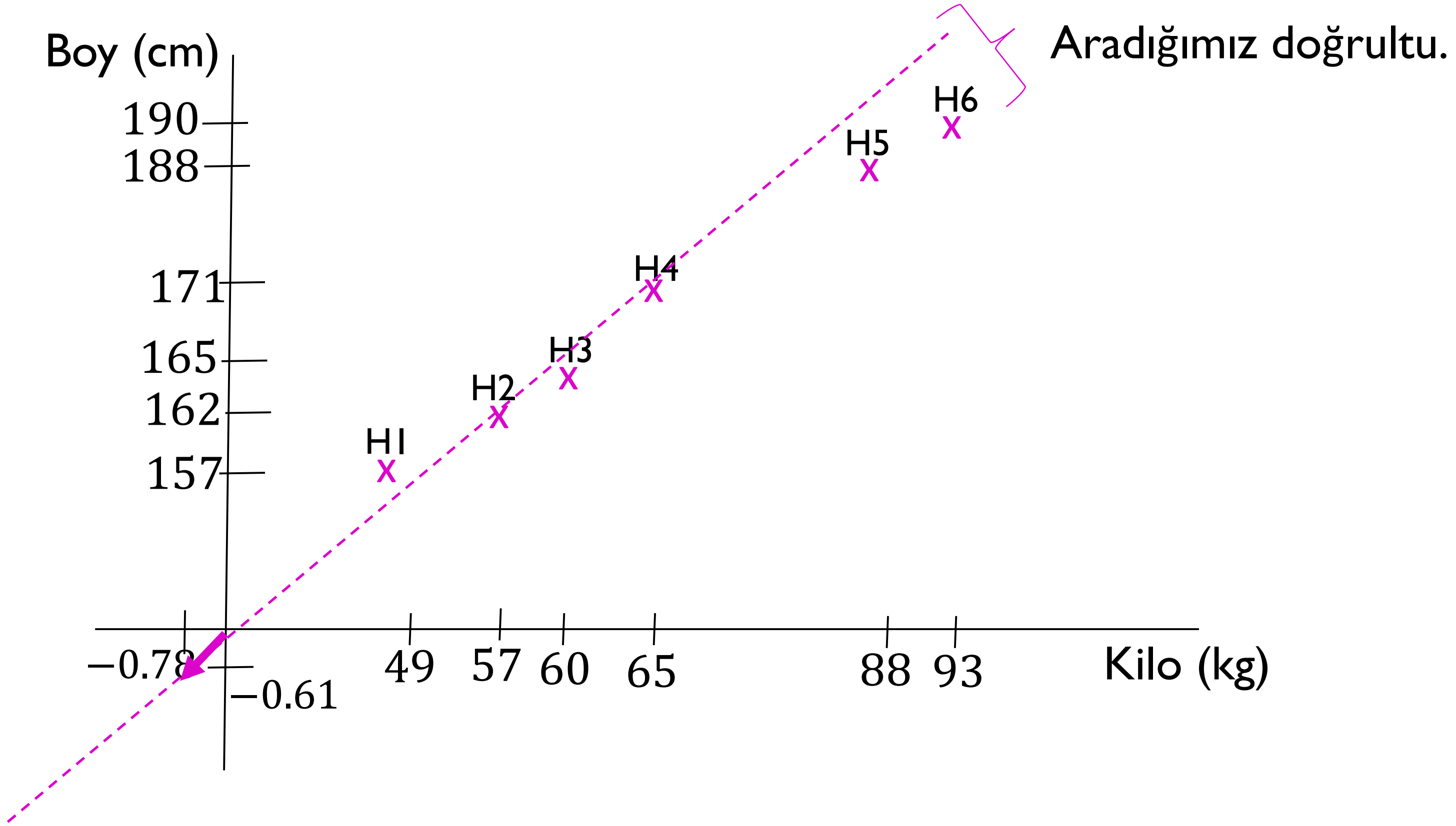
Bir önceki kovayans hesabinda $(57-68)(162-172)$ çarpımını dikkate alirsak iki değer de negatiftir; çarpim dolayisiyla pozitiftir. Bu kişinin hem kilosu ortalamadan (68'den) az, hem de boyu ortalamadan (172'den) azdır. Yani burada ikisi birden azdır; bu iki degisken benzer özellik göstermiştir.

Sonuç olarak elimizdeki veri matrsinine karşılık şöyle bir kovaryans matrisi ortaya çıkmıştır:

$$\Sigma = \begin{bmatrix} \sigma_{kilo}^2 & \sigma_{kilo-boy}^2 \\ \sigma_{boy-kilo}^2 & \sigma_{boy}^2 \end{bmatrix} = \begin{bmatrix} 315 & 244 \\ 244 & 190 \end{bmatrix}$$

(Kovaryans matrisler genelde Σ ile gösterilir.)

Kovaryans matrisin sırasiyla birinci ve ikinci özdeğerleri:505 ve 0.75'tir. İki boyutlu veriyi üzerinde temsil edecegimiz vektör 505'e karşılık gelen özvektördür. Daha önce gördüğümüz şekilde $(\Sigma - 505 \cdot I_2)x = \mathbf{0}$ eşitliğini sağlayan özvektörü hesaplarsak $t\begin{bmatrix} -0.78 \\ -0.61 \end{bmatrix}$ vektörlerini bulabiliriz. $t = 1$ için aradığımız vektör $\begin{bmatrix} -0.78 \\ -0.61 \end{bmatrix}$ vektörüdür. Bunu gösterelim:



H4

Bu doğrultu üzerine H1, H2, …,H6 noktalarını transfer/project edersek:

Artik H1, H2, …, H6 hastalari tek bir boyutta (doğrultuda) gösterilebilir hale geldi. Üstelik İki boyutlu orijinal uzayda kim kime yakinsa yada uzaksa bu tek boyutlu uzayda da bu yakinlik uzaklik korundu. Örneğin eskisinde olduğu gibi transfer sonrasinda da H5 ve H6 birbirine yakin; H6 ve H1 birbirine uzaktır.

Son olarak iki boyutta gösterilen bu hastalari $\begin{bmatrix} -0.78 \\ -0.61 \end{bmatrix}$ vektörü ile iç çarparak onlari tek boyuta indirgeyelim; yani bu doğru üzerindeki saytisal degerlerini bulalım.

H1 için: $< (49,157), (-0.78, -0.61) > = -135$

H2 için: $< (57,162), (-0.78, -0.61) > = -143$

H3 için: $< (60,165), (-0.78, -0.61) > = -148$

H4 için: $< (65,171), (-0.78, -0.61) > = -156$

H5 için: $< (88,188), (-0.78, -0.61) > = -184$

H6 için: $< (93,190), (-0.78, -0.61) > = -190$

Yeni veri matrisi:

Hasta No	?
H1	-135
H2	-143
H3	-148
H4	-156
H5	-184
H6	-190

PCA'ya yapılan en büyük eleştirilerden biri yeni oluşan uzayın bir anlamsal karşılığının olmamasıdır; yani bu uzayın boyutlarının yorumlanamamasıdır. Bize verilen ilk orijinal veri matrsinde birinci boyut hastalarin kilosuna, ikinci boyut ise boylarina karsilik geliyordu. Fakat ortaya çıkan yeni uzayın (tek) boyutunun bir anlamı yoktur. O yüzden yeni veri matrisinde kolonun bir adı yoktur, ? Olarak bırakılmıştır.