Norm ve uzaklık

Dr. Öğr. Üyesi Işık İlber Sırmatel

T.C. Trakya Üniversitesi
Mühendislik Fakültesi
Elektrik - Elektronik Mühendisliği Bölümü
Kontrol Anabilim Dalı

Kaynak (source)

Lecture Slides for Introduction to Applied Linear Algebra: Vectors, Matrices, and Least Squares.
Stephen Boyd, Lieven Vandenberghe

Konu listesi

1. Norm
2. Uzaklık
3. Standart sapma
4. Açı

Bölüm 1

Norm

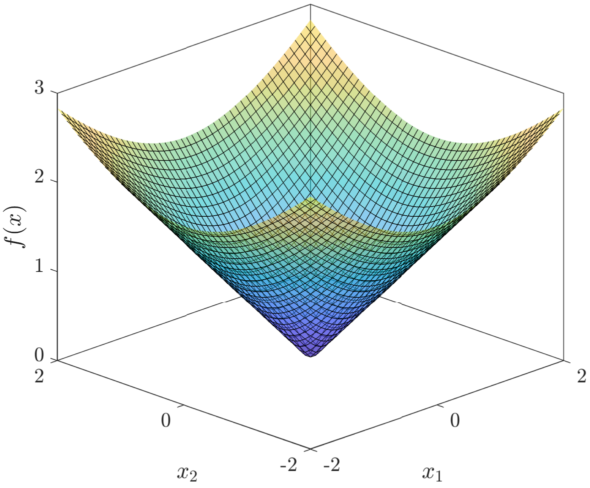
Norm

- norm, bir vektörün büyüklüğünü ölçmek için kullanılır
- bir n-vektörün Öklit normu (veya kısaca normu)

  $$\|x\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} = \sqrt{x^T x}$$

  şeklinde tanımlanır
- Öklit normu $n = 1$ için mutlak değere dönüşür
- Öklit normuna 2-normu da denir (başka isimleri de vardır, örneğin: karesel norm, l^2-normu). diğer normlarla karışmasın diye Öklit normu genellikle $\|x\|_2$ şeklinde gösterilir. biz bu derste sadece Öklit normunu kullanacağımızdan bu normu kısaca $\|x\|$ şeklinde göstereceğiz

$\mathbb{R}^2$ için Öklit normu



$$f(x) = \|x\| \qquad x \in \mathbb{R}^2$$

Normlar (ek bilgi)

Öklit normundan başka vektör normları da vardır, örneğin:

- 1-normu (veya, Manhattan normu)

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|$$

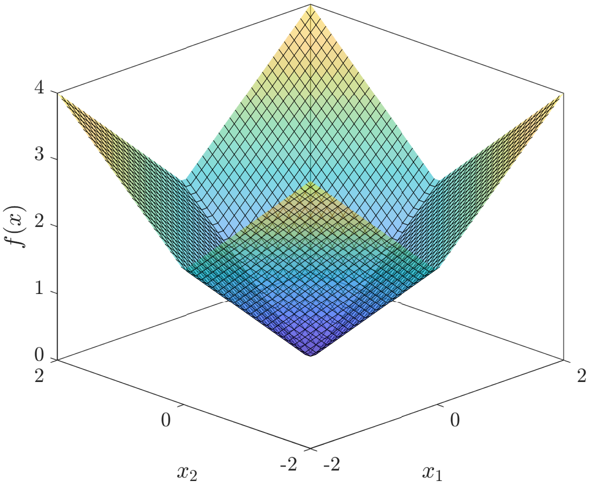
- sonsuz normu

$$\|x\|_\infty = \max(|x_1|, |x_2|, \ldots, |x_n|)$$

- genel olarak: p-normu (p gerçel (*real*) sayı; $p \geq 1$)

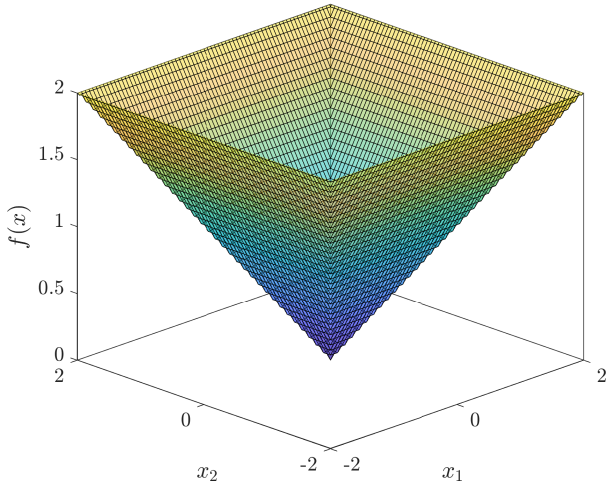
$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}$$

$\mathbb{R}^2$ için 1-normu (ek bilgi)



$$f(x) = \|x\|_1 \qquad x \in \mathbb{R}^2$$

$\mathbb{R}^2$ için sonsuz normu (ek bilgi)



$$f(x) = \|x\|_\infty \qquad x \in \mathbb{R}^2$$

Normun özellikleri

her n-vektör x ve y ile her skaler β için

- homojenlik: $\|\beta x\| = |x|\|x\|$
- üçgen eşitsizliği: $\|x + y\| \leq \|x\| + \|y\|$
- negatif olmama: $\|x\| \geq 0$
- tanımlılık: ancak $x = 0$ ise $\|x\| = 0$

özelliklerine sahip bir $f(x) = \|x\|$ fonksiyonuna norm denir ($f : \mathbb{R}^n \to \mathbb{R}$)

üçgen eşitsizliği haricindekilerin doğru olduğunu göstermek kolaydır; üçgen eşitsizliğinin doğru olduğunu daha sonra göstereceğiz

RMS değeri

- n-vektör x'in ortalama-karesel (*mean-square*) değeri:

$$\frac{x_1^2 + x_2^2 + \cdots + x_n^2}{n} = \frac{\|x\|^2}{n}$$

- n-vektör x'in karekök-ortalama-karesel (*root-mean-square*, RMS) değeri:

$$\mathbf{rms}(x) = \sqrt{\frac{x_1^2 + x_2^2 + \cdots + x_n^2}{n}} = \frac{\|x\|}{\sqrt{n}}$$

- $\mathbf{rms}(x)$, $|x_i|$'nin tipik değerini verir
- örnek: $\mathbf{rms}(\mathbf{1}) = 1$ (n'den bağımsız şekilde)
- RMS değeri farklı uzunluktaki vektörlerin büyüklüklerini karşılaştırmak için kullanışlıdır

Blok vektörlerin normu

- a, b ve c vektörlerini ele alalım. bu vektörler için aşağıdaki ifade geçerlidir

$$\left\| \begin{bmatrix} a \\ b \\ c \end{bmatrix} \right\|^2 = a^T a + b^T b + c^T c = \|a\|^2 + \|b\|^2 + \|c\|^2$$

- bu ifadeden aşağıdaki ifade elde edilir

$$\left\| \begin{bmatrix} a \\ b \\ c \end{bmatrix} \right\| = \sqrt{a^T a + b^T b + c^T c} = \left\| \begin{bmatrix} \|a\| \\ \|b\| \\ \|c\| \end{bmatrix} \right\|$$

(sağ taraftaki ifadeyi doğru anladığınızdan emin olun)

- bu fikirleri daha sonraki kısımlarda kullanacağız

Chebyshev eşitsizliği

- $|x_1|, |x_2|, \ldots, |x_n|$ ile verilen sayılardan k adedinin a'ya eşit veya a'dan büyük olduğunu farz edelim
- o halde $x_1^2, x_2^2, \ldots, x_n^2$ ile verilen sayılardan k adedi a^2'den büyük olacaktır
- dolayısıyla $\|x\|^2 = x_1^2 + x_2^2 + \cdots x_n^2 \geq ka^2$ olur
- dolayısıyla $k \leq \|x\|^2/a^2$ ifadesi geçerlidir
- $|x_i| \geq a$'yı sağlayan x_i'lerin sayısı $\|x\|^2/a^2$'dan fazla değildir
- buna Chebyshev eşitsizliği denir
- RMS değeri cinsinden yazarsak:

$|x_i| \geq a$'yı sağlayan x_i'lerin oranı $\left(\frac{\mathbf{rms}(x)}{a}\right)^2$'dan fazla değildir

- örnek: elemanların en fazla $\%4$'ü $|x_i| \geq 5\ \mathbf{rms}(x)$ ifadesini sağlayabilir

Chebyshev eşitsizliği

örnek:

$$x = \begin{bmatrix} 1.6 \\ 5.5 \\ -6.8 \\ 2.6 \\ 1 \\ -3.9 \\ -1.3 \\ 1 \\ 10.7 \\ 8.3 \\ -4 \\ 9.1 \\ 2.2 \\ -0.2 \\ 2.1 \\ -0.6 \\ -0.4 \\ 4.5 \\ 4.2 \\ 4.3 \end{bmatrix}$$

$$a = 6$$

$|x_i| \geq a$, $k = 4$ adet sayı için sağlanıyor

$$k \leq \frac{\|x\|^2}{a^2} = 12.5858$$

k'nin üst sınırı $= 12.5858$

RMS değeri cinsinden:

$$\left(\frac{\mathbf{rms}(x)}{a}\right)^2 = 0.6293$$

x'in elemanlarından en fazla %62.93'ü

$|x|_i \geq 6$'yı sağlar

Bölüm 2

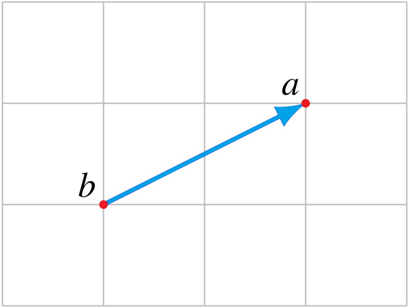
Uzaklık

Uzaklık

- iki n-vektör a ve b arasındaki (Öklit) uzaklık

$$\mathbf{dist}(a, b) = \|a - b\|$$

- $n = 1, 2, 3$ için sıradan (fiziksel) uzaklığı ifade eder



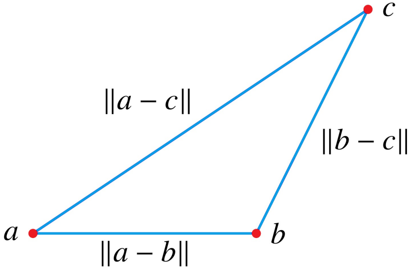
- $\mathbf{rms}(a - b)$, a ile b arasındaki RMS sapmadır

Üçgen eşitsizliği

- köşenoktaları (*vertex*) a, b ve c'de olan bir üçgen ele alalım
- kenar uzunlukları $\|a-b\|$, $\|b-c\|$, $\|a-c\|$
- üçgen eşitsizliği ($\|x+y\| \leq \|x\| + \|y\|$) kullanılarak

$$\|a-c\| = \|(a-b)+(b-c)\| \leq \|a-b\| + \|b-c\|$$

ifadesi yazılabilir, yani üçüncü kenarın uzunluğu diğer iki kenarın uzunluklarının toplamından fazla değildir



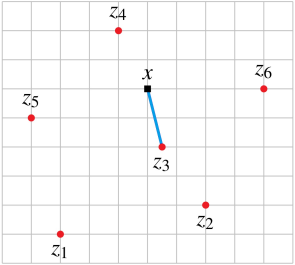
Öznitelik uzaklığı ve en yakın komşular

- x ve y iki ögenin öznitelik vektörleri olsun. bu durumda $\|x - y\|$'ye öznitelik uzaklığı (*feature distance*) denir

- $z_1, z_2, \ldots, z_n$ bir grup vektör olsun. $i = 1, \ldots, m$ için

$$\|x - z_j\| \leq \|x - z_i\|$$

sağlanıyorsa z_j'ye x'in en yakın komşusu (*nearest neighbor*) denir



- bu basit fikirler çok yaygın şekilde kullanılır

Bölüm 3

Standart sapma

Standart sapma

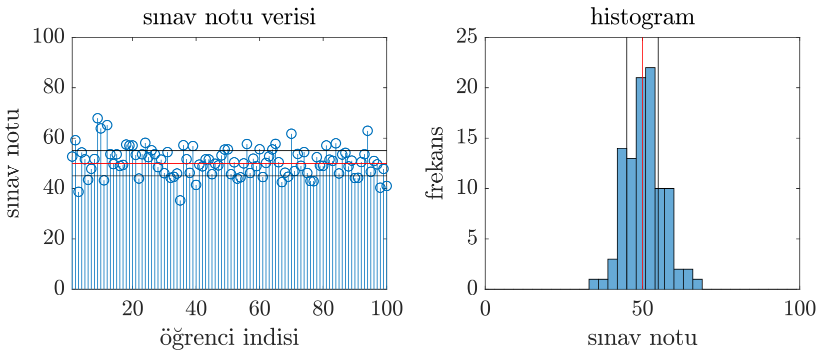
- n-vektör x'in ortalaması (*mean*): $\mathbf{avg}(x) = \mathbf{1}^T x/n$
- x'in ortalamadan arındırılmış (*de-meaned*) hali: $\tilde{x} = x - \mathbf{avg}(x)\mathbf{1}$ ($\mathbf{avg}(\tilde{x}) = 0$ olur)
- standart sapma (*standard deviation*)

$$\mathbf{std}(x) = \mathbf{rms}(\tilde{x}) = \frac{\|x - (\mathbf{1}^T x/n)\mathbf{1}\|}{\sqrt{n}}$$

- $\mathbf{std}(x)$, x_i'nin $\mathbf{avg}(x)$'ten gösterdiği farklılıkların tipik miktarını verir
- ancak $x = \alpha\mathbf{1}$ (bazı α için) ise $\mathbf{std}(x) = 0$ olur
- ortalama ve standart sapma için yaygın olarak Yunan harfleri μ ve σ kullanılır
- RMS değeri ile ortalama ve standart sapma arası bağıntı:

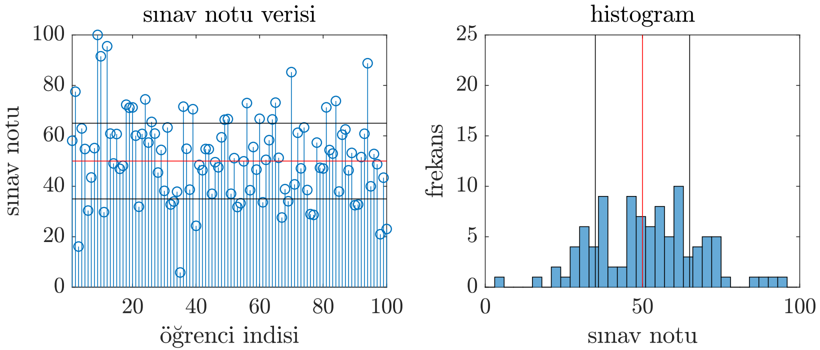
$$\mathbf{rms}(x)^2 = \mathbf{avg}(x)^2 + \mathbf{std}(x)^2$$

Standart sapma: Örnek 1



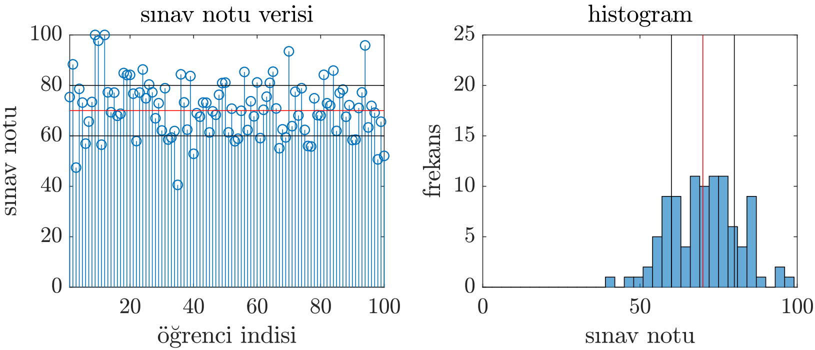
$$\mathbf{avg}(x) = 50 \qquad \mathbf{std}(x) = 5$$

Standart sapma: Örnek 2



$$\mathbf{avg}(x) = 50 \qquad \mathbf{std}(x) = 15$$

Standart sapma: Örnek 3

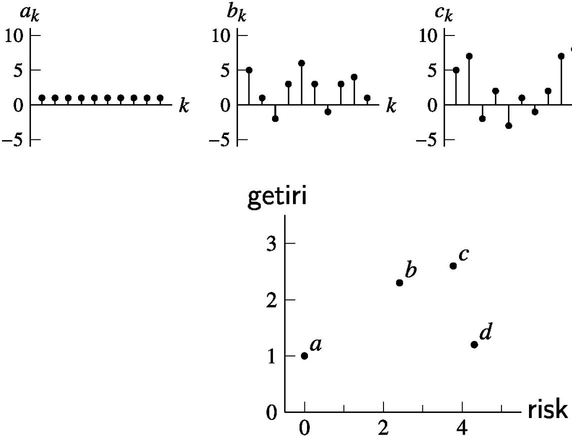


$$\mathbf{avg}(x) = 70 \qquad \mathbf{std}(x) = 10$$

Örnek: Ortalama getiri ve risk

- x bir yatırım (*investment*) (veya değerli varlık (*asset*)) için bir dönem içindeki getirilerin zaman serisi
- $\mathbf{avg}(x)$ dönem için ortalama getiri ($\mathbf{avg}(x)$'e genellikle kısaca getiri denir)
- $\mathbf{std}(x)$ getirinin dönem boyunca ne kadar değişkenlik gösterdiğinin ölçüsüdür. $\mathbf{std}(x)$'e risk denir
- (farklı getiri zaman serileri olan) birden çok yatırım, getiri ($\mathbf{avg}(x)$) ve risk ($\mathbf{std}(x)$) cinsinden karşılaştırılır
- bu karşılaştırma genellikle bir risk-getiri grafiği ile ortaya konur

Örnek: Risk-getiri grafiği



Standart sapma için Chebyshev eşitsizliği

- x n-vektör (ortalaması $\mathbf{avg}(x)$, standart sapması $\mathbf{std}(x)$)
- yaklaşık bir fikir: x'in çoğu elemanı ortalamadan çok uzakta değildir
- Chebyshev eşitsizliğinden, x'in

$$|x_i - \mathbf{avg}(x)| \geq \alpha\ \mathbf{std}(x)$$

  şartını sağlayan elemanlarının oranı $1/\alpha^2$'den fazla değildir ($\alpha > 1$ için)
- 8 ortalama ve 3 standart sapmalı getiri zaman serisi için, kayıp (yani, $x_i \leq 0$) yaşanan dönemler bütün dönemlerin $\%14.1$'inden ($(3/8)^2 = \%14.1$) fazla olmaz

Standart sapma için Chebyshev eşitsizliği

örnek:

$$x = \begin{bmatrix} 1.6 \\ 5.5 \\ -6.8 \\ 2.6 \\ 1 \\ -3.9 \\ -1.3 \\ 1 \\ 10.7 \\ 8.3 \\ -4 \\ 9.1 \\ 2.2 \\ -0.2 \\ 2.1 \\ -0.6 \\ -0.4 \\ 4.5 \\ 4.2 \\ 4.3 \end{bmatrix}$$

$$a = 6$$

$|x_i| \geq a$, $k = 4$ adet sayı için sağlanıyor

$$k \leq \frac{\|x\|^2}{a^2} = 12.5858$$

k'nin üst sınırı $= 12.5858$

RMS değeri cinsinden:

$$\left(\frac{\mathbf{rms}(x)}{a}\right)^2 = 0.6293$$

x'in elemanlarından en fazla %62.93'ü

$|x|_i \geq 6$'yı sağlar

Bölüm 4

Açı

Cauchy-Schwarz eşitsizliği

- iki n-vektör a ve b için Cauchy-Schwarz eşitsizliği

$$|a^T b| \leq \|a\| \|b\|$$

- açık şekilde yazılırsa:

$$|a_1 b_1 + \cdots + a_n b_n| \leq \left(a_1^2 + \cdots + a_n^2\right)^{1/2} \left(b_1^2 + \cdots + b_n^2\right)^{1/2}$$

- buradan üçgen eşitsizliğini gösterebiliriz:

$$\begin{aligned} \|a + b\|^2 &= \|a\|^2 + 2a^T b + \|b\|^2 \\ &\leq \|a\|^2 + 2\|a\| \|b\| + \|b\|^2 \\ &= (\|a\| + \|b\|)^2 \end{aligned}$$

Cauchy-Schwarz eşitsizliğinin türetilmesi

- a veya b 0 ise eşitsizliğin doğru olduğu açıktır
- $\alpha = \|a\|$ ile $\beta = \|b\|$'nin sıfırdan farklı olduğunu varsayalım
- buradan hareketle aşağıdaki ifadeler yazılabilir:

$$\begin{aligned}
0 &\leq \|\beta a - \alpha b\|^2 \\
&= \|\beta a\|^2 - 2\left(\beta a\right)^T\left(\alpha b\right) + \|\alpha b\|^2 \\
&= \beta^2\|a\|^2 - 2\beta\alpha\left(a^T b\right) + \alpha^2\|b\|^2 \\
&= 2\|a\|^2\|b\|^2 - 2\|a\|\|b\|\left(a^T b\right) \\
\Leftrightarrow \quad & \|a\|\|b\|\left(a^T b\right) \leq \|a\|^2\|b\|^2 \\
\Leftrightarrow \quad & a^T b \leq \|a\|\|b\|
\end{aligned}$$

- aynı prosedür $-a$ ve b'ye uygulanarak Cauchy-Schwarz eşitsizliğinin diğer yarısı elde edilebilir

Açı

- sıfırdan farklı iki vektör a ve b'nin arasındaki açı

$$\angle(a,b) = \arccos\left(\frac{a^T b}{\|a\|\,\|b\|}\right)$$

  şeklinde tanımlanır
- $\angle(a,b) \in [0, \pi]$

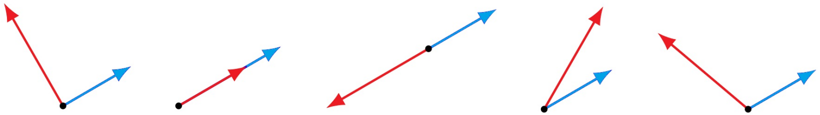
$$a^T b = \|a\|\,\|b\| \cos\left(\angle(a,b)\right)$$

  eşitliğini sağlayan sayıdır
- $n = 2, 3$ için vektörler arası sıradan açıyı ifade eder

Açıların sınıflandırılması

$\theta = \angle(a, b)$

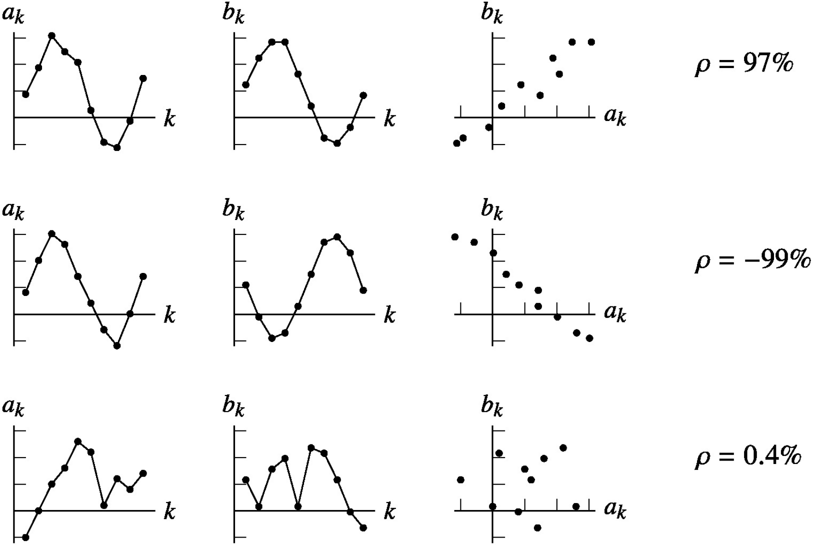
- $\theta = \pi/2 = 90°$: a ve b dikgen (*orthogonal*); $a \perp b$ ($a^T b = 0$) ile gösterilir
- $\theta = 0$: a ve b hizalanmış (*aligned*) ($a^T b = \|a\| \|b\|$)
- $\theta = \pi = 180°$: a ve b ters hizalanmış (anti-aligned) ($a^T b = -\|a\| \|b\|$)
- $\theta \leq \pi/2 = 90°$: a ve b dar açı (*acute angle*) yapar ($a^T b \geq 0$)
- $\theta \geq \pi/2 = 90°$: a ve b geniş açı (*obtuse angle*) yapar ($a^T b \leq 0$)



Korelasyon katsayısı

- a ve b vektörlerini ve ortalamadan arındırılmış halleri
$$\tilde{a} = a - \mathbf{avg}(a)\mathbf{1}, \qquad \tilde{b} = b - \mathbf{avg}(b)\mathbf{1}$$
- a ve b arasındaki korelasyon katsayısı (*correlation coefficient*) ($\tilde{a} \neq 0$, $\tilde{b} \neq 0$)
$$\rho = \frac{\tilde{a}^T\tilde{b}}{\|\tilde{a}\|\|\tilde{b}\|}$$
- $\rho = \cos \angle(\tilde{a}, \tilde{b})$
  - $\rho = 0$: a ve b korelasyonsuz (*uncorrelated*)
  - $\rho > 0.8$ (takriben): a ve b yüksek korelasyonlu (*highly correlated*)
  - $\rho < -0.8$ (takriben): a ve b yüksek ters korelasyonlu (*highly anti-correlated*)
- çok kabaca: a ve b'nin yüksek korelasyonlu olması, a_i ve b_i'nin tipik olarak birlikte ortalamalarının üstünde (veya altında) olması anlamına gelir

Örnekler: Korelasyon katsayısı



Örnekler: Korelasyon

- yüksek korelasyonlu vektörler:
  - yakın konumlardaki yağış zaman serileri
  - aynı sektördeki benzer şirketlerin günlük getirileri
  - yakından ilgili (örneğin, aynı konudaki) belgelerin sözcük sayısı vektörleri
  - ayakkabı ve çorap satışları (farklı konum veya dönemlerde)
- yaklaşık olarak korelasyonsuz vektörler
  - farklı parçaların ses sinyalleri
  - sıcaklık verisi ile borsa endeksi
- (kısmen) ters korelasyonlu vektörler
  - farklı yarı kürelerdeki iki şehirdeki günlük sıcaklıklar