

## Kovaryans ve Korelasyon

Bugun “kovaryans gunu”, bu teknigi kullanarak nihayet bir toplamin varyansini bulabilecegiz, varyans lineer degildir (kiyasla beklenti -expectation- lineerdir). Bu lineer olmama durumu bizi korkutmayacak tabii, sadece yanlis bir sekilde lineerlik uygulamak yerine probleme farkli bir sekilde yaklasmayi ogrenecegiz.

Diger bir acidan, hatta bu ana kullanimlardan biri, kovaryans iki rasgele degiskeni beraber / ayni anda analiz etmemize yarayacak. Iki varyans olacak, ve onlari alakasina bakiyor olacagiz, bu sebeple bu analize *kovaryans* deniyor zaten.

### Tanim

$$\text{Cov}(X, Y) = E((X - E(X))(Y - E(Y))) \quad (1)$$

Burada  $X, Y$  ayni uzayda tanimlanmis herhangi iki rasgele degisken. Ustteki diyor ki bu iki rasgele degisken  $X, Y$ 'in kovaryansi,  $X$ 'ten ortalamasi cikartilmis,  $Y$ 'ten ortalamasi cikartilmis halinin carpilmasi ve tum bu carpimlari ortalamasinin alınmasidir.

Tanim boyle. Simdi bu tanima biraz bakip onun hakkında sezgi / anlayis gelistirmeye ugrasalim. Tanim niye bu sekilde yapilmis, baska bir sekilde degil?

Ilk once esitligin sag tarafindaki bir carpim, bir sey carpi bir baska sey. Bu seylere den biri  $X$  ile digeri  $Y$  ile alakali, onlari carparak ve carpimin bir ozelliginden faydalanarak sunu elde ettik; arti deger carpi arti yine arti degerdir, eski carpi arti eksidir, eksi carpi eksi artidir. Bu sekilde mesela “ayni anda arti” olmak gibi kuvvetli bir baglanti carpimin arti olmasi ile yakalanabilecektir. Ayni durum eksi,eksi de icin gecerli, bu sefer her iki rasgele degisken ayni sekilde negatiftir. Eksi carpim sonucu ise sifirdan az bir degerdir, “kotu korelasyon” olarak alinabilir ve hakikaten de eksi arti carpiminin isareti oldugu icin iki degiskenin ters yonlerde oldugunu gosterir. Demek ki bu arac / numara hakikaten faydalidir.

Unutmayalim, ustteki carpimlardan birisinin buyuklugu  $X$ 'in ortalamasina bagli olan bir diger,  $Y$  ayni sekilde. Simdi  $X, Y$ 'den bir orneklem (sample) aldigimizi dusunelim. Veri setinin her veri noktasini bagimsiz ve ayni sekilde dagilmis (i.i.d) durumda. Yani  $X, Y$  degiskenlerine “gelen”  $x_i, y_i$  ikilileri her  $i$  icin digerlerinden bagimsiz; fakat her ikilinin arasinda bir baglanti var, yani demek ki bu rasgele degiskenlerin baz aldigi dagilimlari bir alakasi var, ya da bu iki degiskenin bir birlesik dagilimi (joint distribution) var.

Not: Eger  $X, Y$  bagimsiz olsaydi, o zaman

$$\text{Cov}(X, Y) = E((X - E(X))E(Y - E(Y)))$$

olarak yazilabilirdi, yani iki beklentinin ayri ayri carpilabildigi durum... Ama biz bu derste bagimsizligin olmadigi durumla ilgileniyoruz..

Korelasyon kelimesinden bahsedelim hemen, bu kelime gunluk konusmada cok

kullaniliyor, ama bu ders baglaminda korelasyon kelimesinin matematiksel bir anlami olacak, onu birazdan, kovaryans uzerinden tanimlayacagiz.

Bazi ilginc noktalar:

Ozellik 1

varyansi nasil tanimlamistik?

$$\text{Var}(X) = E((X - E(X))^2)$$

Bu denklem aslinda

$$\text{Cov}(X, Y) = E((X - E(X))(Y - E(Y)))$$

denkleminde Y yerine X kullandigimizda elde ettigimiz seydir, yani

$$\text{Cov}(X, X) = E((X - E(X))(X - E(X)))$$

$$\text{Cov}(X, X) = E((X - E(X))^2)$$

$$= \text{Var}(X)$$

Yani varyans, bir degiskenin “kendisi ile kovaryansidir”. Ilginc degil mi?

Ozellik 2

$$\text{Cov}(X, Y) = \text{Cov}(Y, X)$$

Ispati kolay herhalde, (1) formülünü uygulamak yeterli.

Teori

$$\text{Cov}(X, Y) = E((X - E(X))(Y - E(Y))) = E(XY) - E(X)E(Y)$$

Ispat

Bu ispat cok kolay, esitligin sol tarafindaki carpimi parantezler uzerinden acarsak, ve beklenti lineer bir operator oldugu icin toplamın terimleri uzerinde ayri ayri uygulanabilir,

$$E(XY) - E(X)E(Y) - E(X)E(Y) + E(X)E(Y)$$

$$= E(XY) - E(X)E(Y)$$

Carpimi uygularken mesela  $E(-X \cdot E(Y))$  gibi bir durum ortaya çıktı, burada  $E(Y)$ 'nin bir sabit olduğunu unutmayalım, çünkü beklenti rasgele degiskene uygulanınca tek bir sayı ortaya çıkartır, ve şu  $E(Y)$  üzerinde bir beklenti daha uygulanınca bu “icerideki” beklenti sabitmiş gibi dışarı çıkartılabilir, yani  $-E(X)E(Y)$ .

Devam edelim,  $E(XY) - E(X)E(Y)$  ifadesini gösterdik, çünkü çoğu zaman bu ifade hesap açısından (1)'den daha uygundur. Ama (1) ifadesi anlatım / sezgisel kavrayış açısından daha uygun, çünkü bu ifade  $X$ 'in ve  $Y$ 'nin kendi ortalamalarına izafi olarak belirtilmiştir, ve akılda canlandırılması daha rahat olabilir. Fakat matematiksel olarak bu iki ifade de aynidir.

İki özellik bulduk bile. Bir özellik daha,

Ozellik 3

$$\text{Cov}(X, c) = 0$$

Bu nereden geldi? (1)'e bakalım,  $Y$  yerine  $c$  koymuş olduk, yani bir sabit. Bu durumda (1)'in  $(Y - E(Y))$  kısmı  $c - E(c) = c - c = 0$  olur [aslında bayagi absürt bir durum], ve bu durumda (1) tamamen sifira dönüşür, sonuç sıfır.

Ozellik 4

$$\text{Cov}(cX, Y) = c \cdot \text{Cov}(X, Y)$$

İspat için alttaki formülde

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$$

$X$  yerine  $cX$  koymak yeterli,  $c$  her iki terimde de dışarı çıkacaktır, ve grubun dışına alınca bu özelliği elde ederiz.

Ozellik 5

$$\text{Cov}(X, Y + Z) = \text{Cov}(X, Y) + \text{Cov}(X, Z)$$

İspat için bir üstteki özellikte yaptığımızın benzerini yaparız.

En son iki özellik oldukça faydalıdır bu arada, onlara ikili-lineerlik (bilinearity) ismi veriliyor. İsim biraz renkli / başarılı bir isim, söylemek istediği şu aslında, bu son iki özellikte sanki bir koordinatı sabit tutup diğeri ile işlem yapmış gibi oluyoruz, yani bir koordinat sabit olunca diğeri “lineer” gibi oluyor; Mesela  $c$ 'nin dışarı çıktığı durumda olduğu gibi, bu özellikte  $Y$ 'ye hiçbir şey olmadı, o degismeden kaldı. Aynı şekilde 5. özellikte  $X$  hiç degismeden esitliğin sağına aktarıldı sanki, sadece “ $Z$  durumu için” yeni bir terim ekledik.

4. ve 5. ozellik cok onemlidir, bunlari bilerseniz bir ton hesabi yapmadan hizlica tureterek hesaplarinizi kolaylastirabilirsiniz.