İST 304 İstatistik Karar Kuramı ve Yöntemleri **Hafta VIII**

Ders 1 : Doğanın Belirsizliğinde Bayes İlkesi II

Dersi anlatan: İhsan Karabulut *Notları yazan: İhsan Karabulut*

Not: *LaTeX ders kalıbı UC Berkeley EECS Bölümü'nündür.*

Uyarı: *Bu ders notları formal yayınların tabi olduğu kanun, yönetmelik, kural ve esaslar dışındadır. Ders dışında herhangi bir şekilde kopya edilmesi, çoğaltılması, yayımlanması yalnızca bu notları hazırlayan ve yazanın iznini gerektirir.*

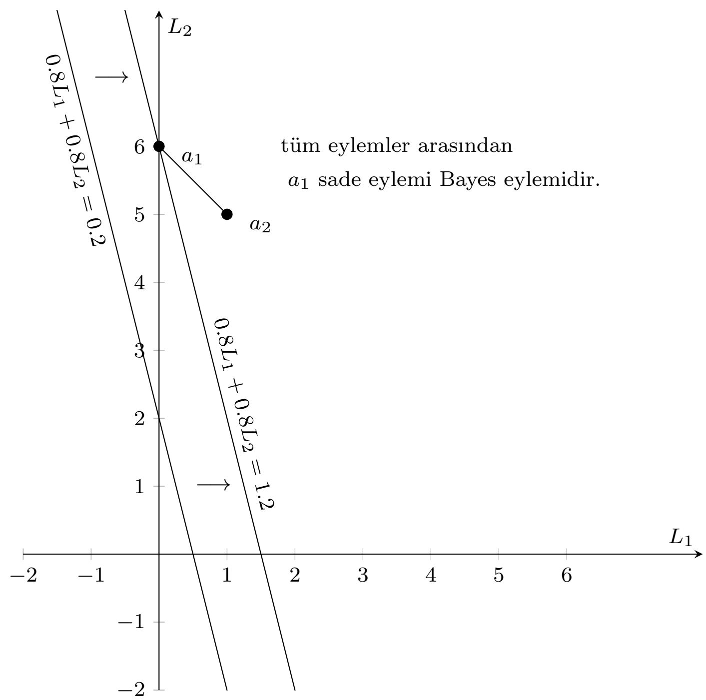
Örnek. (Asansör problemine devam) Herhangi bir karma eylem $p = (p, 1-p)$ için beklenen kayıp $E(l(\theta_1, a)) = L_1 = p.0 + 1.(1-p) = 1-p$, $E(l(\theta_2, a)) = L_2 = 5+p$ olacaktır. Doğanın durumu için önsel dağılım $g(\theta_1) = 0.8$, $g(\theta_2) = 0.2$ olmak üzere herhangi bir karma ya da sade eylem için Bayes kaybı:

$$B(p) = 0.8L_1 + 0.2L_2$$

dır. Verilen bir sabit c değeri için $B(.) = c$ olduğu eylemlerin tümünün kümesi (L_1, L_2) noktalarının düzleminde

$$0.8L_1 + 0.2L_2 = c$$

doğrusu üzerinde yer alacaktır. Söz konusu doğru parçasının eğimi $-0.8/0.2 = -4$ dür. Asansör probleminde iki sade eylemle oluşturulan konveks kabuğun bir doğru parçası olduğu daha önce gösterilmişti. Şekil grafik ile Bayes eyleminin saptanmasını göstermektedir. c sabiti konveks kabuğa doğru parçası konveks kabuğa doğru ötelenirken konveks kabukta değdiği ilk noktayı temsil eden karma ya da sade eylem (ya da eylemler) Bayes eylemi olacaktır. Doğrunun konveks kabuğa doğru ötelenmesi c değeri değiştirilerek yapılır. Doğrunun konveks kabuğa ilk kez değdiği durumdaki c değeri de söz konusu eylemin Bayes kaybı değeridir. Minimaks eyleminin görsel olarak grafikle araştırılmasında olduğu gibi Bayes eyleminin araştırılmasında da doğrunun ve konveks kabuğun birer konveks küme oldukları unutulmamalı; yüksek boyutlarda çözümün varlığı ve sayısal yöntemlerle en iyinin araştırılması bu koşullara bağlıdır.



Şekil 1.1: Asansör probleminde (0.8, 0.2) önsel dağılımı ile tüm sade ve karma eylemler arasından Bayes eyleminin grafik ile belirlenmesi.

$0.8L_1 + 0.2L_2 = c$ doğrusunda c ye deneme amaçlı değerler verilip artırılarak a_1 ve a_2 sade eylemlerine ilişkin $(1, 5)$, $(0, 6)$ noktalarının oluşturduğu konveks kümeye(doğru parçasına) yaklaştırılır. Burada grafiğin faydası bu yaklaştırmayı izleyebilmek ve doğrunun kümenin hangi noktasına değebileceğinin görülerek konumunun saptanmasıdır. Konum saptandıktan sonra analitik olarak Bayes eylemine ait (L_1, L_2) noktası ve Bayes kaybı bulunur. Şekil'de verilen çizime koşut olarak işlemler aşağıdaki sırayı izleyecektir:

1) Sade eylemlere ait (L_1, L_2) beklenen kayıp vektörlerini kullanılarak tüm sade ve karma eylemlere ait (L_1, L_2) noktalarını içeren konveks kabuk oluşturulur. Asansör örneğinde uç noktalarının a_1 ve a_2 sade eylemlerine ait $(L(\theta_1, a_i), L(\theta_2, a_i))$ noktalarının olduğu doğru parçasıdır.

2) Herhangi bir eylemin verilen önsel dağılım altında Bayes kaybı tanımını uygulayarak Bayes kaybı ifade edilir. Verilen örnekte bu $B(a) = B(p) = 0.8L_1 + 0.2L_2$ dır. Eğimi önsel dağılımca belirlenmiş olan bir doğruyu tanımlar. Bu herhangi bir c değerine eşit olacaktır ve $0.8L_1 + 0.2L_2 = c$ daha önce R^2'de bir doğrunun $ax + by = c$ şeklindeki bir ifadesidir. c artırılıp azaltılarak doğru yukarı doğru veya aşağı doğru kendine paralel olarak ötelenebilir.

3) İlk denemede $c = 0.4$ seçilmiştir ve $0.8L_1 + 0.2L_2 = 0.4$ doğrusu Şekil'deki çizimde gösterilmiştir. Doğrunun çizimi iki şekilde de yapılabilir. İlki $0.8L_1 + 0.2L_2 = 0.4$ doğrusunun $L_2 = 2 - 4L_1$ olarak ifade edip L_1'e değerler vererek çizmektir. Diğeri $0.8L_1 + 0.2L_2 = 0.4$ için önce $L_1 = 0$ iken $L_2 = 2$ olduğunu hesaplayıp R^2'de bir doğruyu tanımlamak için gerekli noktalardan birini $(0, 2)$ elde edilir. Sonra $L_2 = 0$ iken $L_1 = 1/2$ olduğunu hesaplayıp R^2'de bir doğruyu tanımlamak için gerekli noktalardan ikincisi $(1/2, 0)$ elde edilir. Bu iki noktadan geçen doğru yine ilk elde edilen doğruyla aynı olacaktır. Bu aşamada ne $c = 0.4$ seçilmesi ve L_1 ve L_2' verilen değerler özeldirler, herhangi bir şekilde belirlenebilirler, daha az sayıda deneme yapmak deneyim gerektirir.

4) $0.8L_1 + 0.2L_2 = 0.4$ doğrusunun yukarı doğru ötelendiğinde ilk kez konveks kabuk üzerindeki a_1 sade eylemini temsil eden $(0, 6)$ noktasına değeceği görülür. Bu nokta ile $0.8L_1 + 0.2L_2 = c$ doğrusunun çakışması doğrunun $L_1 = 0$ ve $L_2 = 6$ noktasının doğru üzerinde yer alması ile ifade edilir: $0.8 \times 0 + 0.2 \times 6 = 1.2$. O halde önsel dağılımca eğimi belirlenen $0.8L_1 + 0.2L_2 = c$ doğruları içinde $0.8L_1 + 0.2L_2 = 1.2$ doğrusu konveks kabuğa ilk kez değen doğrudur.

5) Verilen önsel dağılım altında $(L_1, L_2) = (0, 6)$ noktası ile tanımlanan a_1 sade eylemi Bayes eylemidir ve Bayes kaybı $B(a_1) = 1.2$ dir.

Kayıp fonksiyonu ve minimaks ilkesi kullanılarak belirlenen minimaks eylemi ile ile pişmanlık fonksiyonu kullanılarak elde edilen minimaks eyleminin her zaman aynı olmayacağından söz edilmişti. **Kayıp veya pişmanlık fonksiyonunun Bayes ilkesi ile kullanılması ile belirlenen Bayes**

eylemlerinin her zaman aynı olduğu görülebilir. Verilen $g(\theta)$ önsel dağılım altında bir a_j sade eylemi için kayıp fonksiyonu $l(\theta_i, a_j)$ yerine

$$r(\theta_i, a_j) = l(\theta_i, a_j) - \min_a l(\theta_i, a)$$

pişmanlık fonksiyonu kullanılarak **Bayes pişmanlığı**

$$E(r(\theta, a_j)) = \sum_{i=1}^{m} g(\theta_i) r(\theta_i, a_j)$$

olarak tanımlanabilir. Bu ifadede pişmanlık fonksiyonun tanımı yerine konularak

$$E(r(\theta, a_j)) = E(l(\theta, a_j)) - \sum_{i=1}^{m} g(\theta_i) \min_a l(\theta_i, a)$$

elde edilir. Burada $\sum_{i=1}^{m} g(\theta_i) \min_a l(\theta_i, a)$ terimi a_j'ye değil her θ_i için $\min_a l(\theta_i, a)$ değerlerine bağlıdır ve sabittir. Sade bir eylemin Bayes pişmanlığı $E(r(\theta, a_j))$ ile sade bir eylemin Bayes kaybı arasındaki fark $\sum_{i=1}^{m} g(\theta_i) \min_a l(\theta_i, a)$kadar olacaktır ve a_j'den bağımsız sabittir. Böylelikle kayıp fonksiyonuna göre en küçük kaybı sağlayan eylem pişmanlık fonksiyonuna göre de en küçük pişmanlığı sağlar; x_i, i eyleminin kayıp değerini ve $x_i + c$ i eyleminin pişmanlık değerini göstermek üzere $\min_{x_i}\{x_1, x_2, \ldots, x_m\} = x_r$, $\min_{x_i+c}\{x_1 + c, x_2 + c, \ldots, x_m + c\} = x_r + c$ dir.

Not. Bayes pişmanlığı için yeni bir notasyon kullanılmadı, yeri geldiğinde Bayes pişmanlığı vurgusu yapılacaktır.

Örnek. Aşağıdaki kayıp fonksiyonuna sahip karar verme problemi dikkate alınsın.

		$l(\theta_i,a_j)$		
	a_1	a_2	a_3	$\min_a l(\theta_i,a)$
θ_1	2	5	3	2
θ_2	3	1	5	1

Burada pişmanlıklar her a sade eylemi ve iki doğa durumunda $r(\theta_1, a) = l\,(\theta_1, a) - 2$ ve $r(\theta_2, a) =$

$l(\theta_2, a) - 1$ olarak hesaplanacaklardır. $g(\theta)$ önsel olasılık dağılımına göre Bayes pişmanlığı :

$$\begin{aligned} E(r(\theta, a_1)) &= \sum_{i=1}^{2} g(\theta_i) r(\theta_i, a_1) \\ &= g(\theta_1)(l(\theta_1, a_1) - 2) + g(\theta_2)(l(\theta_2, a_1) - 1) \\ &= B(a_1) - (2g(\theta_1) + g(\theta_2)) \end{aligned}$$

$E(r(\theta, a_2)) = B(a_2) - (2g(\theta_1) + g(\theta_2))$ ve $E(r(\theta, a_3)) = B(a_3) - (2g(\theta_1) + g(\theta_2))$ benzer olarak elde edilir. $B(a_i), i = 1, 2, 3$ arasından hangi eylem verilen önsel dağılımla Bayes eylemi ise pişmanlıklara dayalı Bayes eylemi de bu eylem olacaktır.

Örnek. Karadeniz'de bir yerleşim yerinde havanın iki durumu vardır. θ_1 havanın açık ve güneşli olmasını, θ_2 havanın yağmurlu olmasını göstersin. Buraları gezmekte olan Turist Ömer'in ise hava durumu karşısında üç sade eylemi vardır:

a_1: Açık ve güneşli havaya uygun giyim,a_2: Yağışlı hava durumunda sadece üstünün ıslanmamasını sağlayabilecek bir yağmurluk giyinmek ve a_3: Yağmurluk, bot, yağmur şapkası ve şemsiye alarak tam olarak yağmurlu havaya hazırlıklı giyinmek. Turist Ömer'in kayıp fonksiyonu da aşağıdaki gibidir:

		$l(\theta_i,a_j)$	
	a_1	a_2	a_3
θ_1	0	1	3
θ_2	5	3	2

Bulunulan yörede havanın durumuna ilişkin önsel dağılım bu güne kadar deneyim ve gözlem sonucu olarak $P(\theta = \theta_1) = 1/3$, $P(\theta = \theta_2) = 2/3$ olduğu düşünülmektedir. Söz konusu karar verme problemi için tüm sade ve karma eylemler arasından grafik kullanılarak Bayes eylemi belirlenmek istensin.

θ_i doğa durumunda herhangi bir a karma eylemi için beklenen kayıp $L_i = L(\theta_i, a)$ olmak üzere

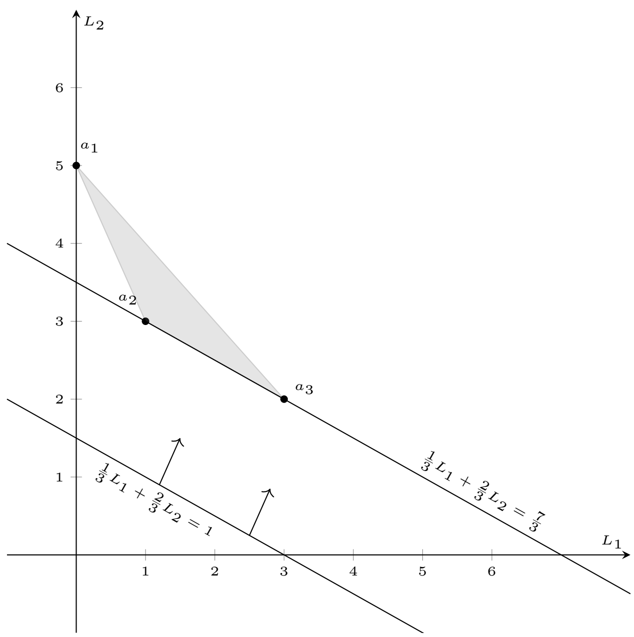
$$L_1 = E(l(\theta_1, a)) = 0.p_1 + p_2 + 3p_3$$

$$L_2 = E(l(\theta_2, a)) = 5p_1 + 3p_2 + 2p_3$$

olacaktır. Bilindiği gibi $a = a_1$sade eylemi $(1,0,0)$ dağılımı , $a = a_2$ eylemi $(0,1,0)$ dağılımı ve $a = a_3$ eylemi $(0,0,1)$ dağılımı ile ifade eder. Herhangi bir karma eylem $p = (p_1, p_2, p_3)$ olasılık dağılımıyla ve verilen önsel dağılım için bu eylemin Bayes kaybı

$$B(p) = \frac{1}{3}L_1 + \frac{2}{3}L_2$$

olarak ifade edilecektir. Bütün sade ve karma eylemlere ait (L_1, L_2) noktalarından oluşan en küçük konveks küme üç sade eyleme ilişkin (L_1, L_2) noktalarının konveks kabuk Şekil'de olduğu gibidir.



Şekil 1.2: Turist Ömer'in karar verme probleminde önsel dağılım $(\frac{1}{3}, \frac{2}{3})$ olduğunda Bayes eyleminin grafikle belirlenmesi.

Grafikten de görüldüğü gibi $a_2$veya a_3 sade eylemleri ile bu eylemlerin herhangi bir karması verilen önsel dağılıma göre birer Bayes eylemidirler. Bayes eylemlerinin kaybı da $7/3$ dır. Örneğin $a_2$ve a_3 eylemlerine karşılık gelen noktaları birleştiren doğru parçası üzerinde temsil edilen $[a_1, a_2, a_3]_{(0,2/5,3/5)}$ eylemi bir Bayes eylemidir. Bu eylemler içinden bir sade eylem seçilirse uygulama da kolaylaşır;

Bayes eylemlerinin uygulama kolaylıklarından birisi de budur.

Bu durum şöyle açıklanabilir. Bayes eylemini belirleme amacıyla çizilen doğru pek çok önsel dağılım için konveks kümeye ilk kez bir sade eylemi temsil eden **köşeye** değecek veya göreli olarak az da olsa bu doğru bazı önsel dağılımlar için konveks kümenin bir sınırını oluşturan doğru parçasına parelel olacak ve tam olarak bu doğru parçasına yapışacaktır. Dolayısıyla bu doğru parçasının uçlarında yer alan sade eylemler de tıpkı bunların herhangi bir karması gibi bir Bayes eylemi olacaktır, aynı Bayes kaybına sahip olacaktır.

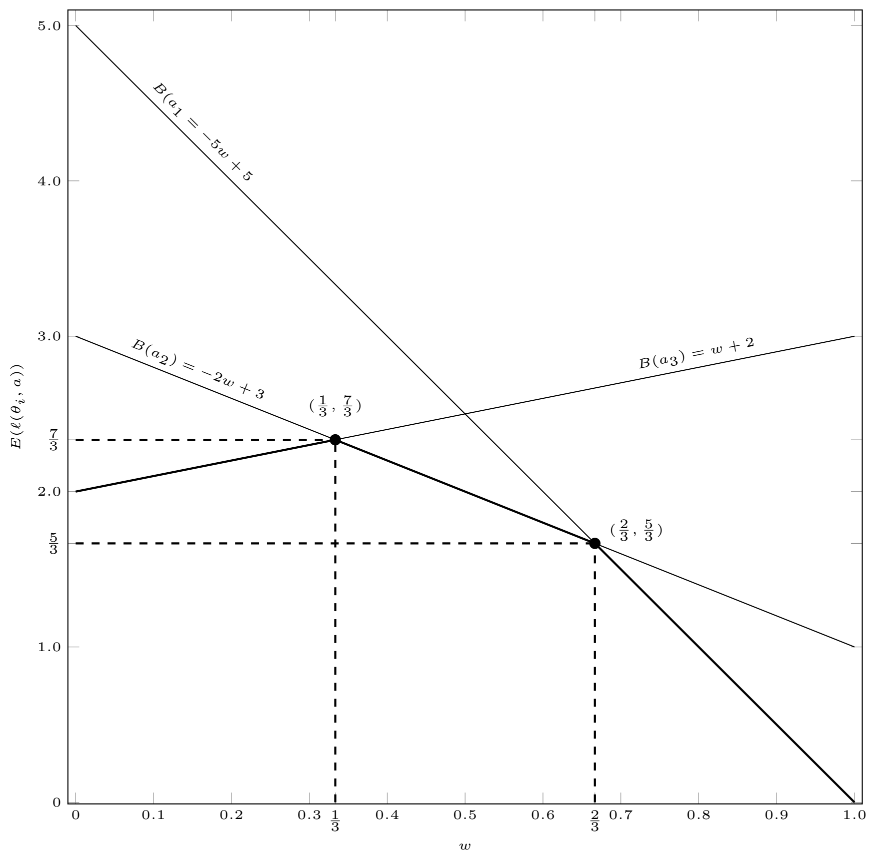
Doğanın sadece iki durumunun olduğu Turist Ömer probleminde verilen herhangi bir $g(\theta_1) = w,\ g(\theta_2) = 1 - w$ önsel dağılım için yalnızca sade eylemlerin Bayes kayıpları göz önüne alınarak bunlardan hangisinin (ya da hangilerinin) Bayes eylemi olabilecekleri yine bir grafik ile saptanabilir. Sade eylemlerin Bayes kayıpları:

$$
\begin{aligned}
B(a_1) &= \sum_{i=1}^{2} g(\theta_i) l(\theta_i, a_1) \\
&= w.0 + (1-w).5 \\
&= -5w + 5
\end{aligned}
$$

ve benzer olarak $B(a_2) = -2w + 3$, $B(a_3) = w + 2$ bulunur. Her birinin Bayes kaybı önsel dağılımın (doğrusal)bir fonksiyonudur; bu fonksiyonlar aynı düzlemde aşağıda Şekil'de olduğu gibi gösterilebilirler.

Daha önce de değinildiği gibi Şekil'de doğrusal fonksiyonlarla oluşturulan ve kalın çizgilerle belirlenmiş bir fonksiyon zarfı görülmektedir. Kalın çizgiler üzerindeki $(w, B(a))$ noktalar kümesi iki doğa durumu ile tanımlanan karar probleminde verilen $(w, 1-w)$ önsel dağılım altında $B(a)$ Bayes kaybı en küçük olan a eylemlerini de belirtmektedir.

Şekilden $w \in [0, 1/3]$ olduğunda a_3, $w \in [1/3, 2/3]$ olduğunda a_2 ve $w \in [2/3, 1]$ olduğunda ise a_1 eylemi Bayes eylemi olacaktır. örneğin, $w = 1/3$ olduğunda karar verici a_2 ve a_3 sade eylemlerinden



Şekil 1.3: Turist ömer probleminde sade eylemlerin değişik önsel dağılımlara göre Bayes kayıpları.

birini belirleyebilir. Bunların yanında bu sade eylemleri rasgeleleştirdiği bir başka eyleme de karar verebilir.

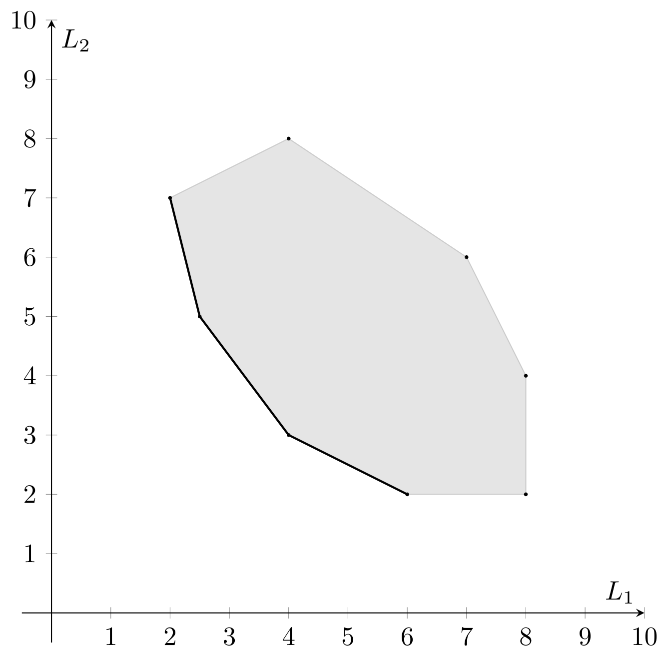
Buraya kadar sadece iki doğa durumunun dikkate alındığı örnekler üzerinde duruldu. Buraya kadar sunulan anlayış kullanılarak herhangi k tane doğa durumunun yer aldığı bir karar problem ele alınabilir, ancak problemin geometrik olarak çözümlenmesi güçleşecektir. Örneğin $k = 3$ olduğunda Bayes kararının belirlenmesinde kullanılan doğru yerine bir düzlem ayırt edici düzlem olarak kullanılacak, minimaks kararının belirlenmesinde kullandığımız karesel konveks küme yerini bir küp alacaktır. $k > 3$ olduğunda da yapılabilirse boyut indirgenerek problem görselleştirilip çözülmeye çalışılacaktır; geometrik çözüm arayışının sınırlı olduğu görülecektir. $k \geq 3$ olduğunda karar problemlerinin çözümlenmesi ayırdedici düzlem teoreminin koşulları altında -çözümlerin var olduğu garantisi altında - analitik olarak, çoğu durumda sayısal çözüm yöntemleri kullanılarak yapılabilmektedir.

İST 304 İstatistik Karar Kuramı ve Yöntemleri **Hafta VIII**

Ders 2 : Kabul Edilebilirlik ve Karar Verme İlkeleri I

Dersi anlatan: İhsan Karabulut *Notları yazan: İhsan Karabulut*

İki doğa durumunun olduğu karar verme problemlerinde bütün sade ve karma eylemlerin kümesinin bir konveks küme olduğu biliniyor. Minimaks kararın veya Bayes kararının belirlenmesi problemlerinde bu kümenin sade yada karma eylemleri temsil eden kimi elemanlarının hiçbir zaman, çözüme konu olamayacağı ve bu nedenle kullanılamayacağı dikkati çekmiş olmalıdır. Nedeni de açık gibi görünüyor: Bütün doğa durumlarında kayıpları bu eylemlerin kayıplarından daha küçük kayba sahip eylemler varken bu eylemlerin kullanılması kabul görmez.



Şekil 2.4: İki doğa durumunda sade eylemlerle oluşturulan tipik bir konveks kabukta yer alan kabul edilebilir eylemler koyu çizgiler üzerinde (L_1, L_2) noktalarıyla temsil edilen eylemlerdir.

Aşağıdaki tanımlamalar bu konuda verilecek kavramları işevuruklaştırmak için kullanılacaktır. Tanımlar

k tane doğa durumunun yer aldığı bir karar problemi dikkate alınarak yapılacak, a, a^* eylem kümesinde yer alan sade ya da karma herhangi iki eylemi gösterecektir.

Tanım. a, a^* eylem kümesinde yer alan herhangi iki eylem olsun. $L(\theta_1, a) \leq L(\theta_1, a^*)$, $L(\theta_2, a) \leq L(\theta_2, a^*)$, $L(\theta_3, a) \leq L(\theta_3, a^*)$, $\ldots$, $L(\theta_k, a) \leq L(\theta_k, a^*)$ ise a **eylemi en az** a^* **eylemi kadar iyidir** (at least as good as) denilir.

Tanım. $L(\theta_1, a) = L(\theta_1, a^*)$, $L(\theta_2, a) = L(\theta_2, a^*)$, $L(\theta_3, a) = L(\theta_3, a^*)$, $\ldots$, $L(\theta_k, a) = L(\theta_k, a^*)$ ise a **eylemi** a^* **eylemine denktir** denilir.

Bu tanımlama ile birlikte bir hatırlatma uygun olacaktır. Eğer iki eylem denk iseler bu eylemler tanımlama olarak farklı olsalar da R^k de aynı nokta ile temsil edileceklerdir.

Tanım. a en az a^* eylemi kadar iyi ve a^*eylemine denk değilse a **eylemi** a^* **eylemine baskındır** denilir.

Tanım. a eylemine baskın bir eylem yoksa a **kabul edilebilir** bir eylemdir denilir.

Kabul edilebilir sade yada karma eylemler bütün eylemlerin kümesi olan konveks kabuğun sınır noktalarının bir parçasıdır. Sınır noktalarının bu parçası **sınırın kabul edilebilir parçası** olarak adlandırılacaktır.

Karar verici hangi karar verme ilkesini kullanırsa kullansın vereceği karara konu olan eylem kabul edilebilir eylemlerden biri olaması beklenir. Minimaks ve Bayes ilkeleri kullanılarak elde edilen eylemler bu özelliğe sahiptirler. Karar verici kararını sadece kabul edilebilir eylemlerle sınırlayarak da verebilir. Bu çoğu kez makuldür, fakat bazen karar verme probleminin yapısından gelen sorunlu (pathological) durumlarla karşılaşmak nadiren de olsa söz konusudur. Konveks kümenin sınır noktaları bu kümeye dahil olmaz ise aşağıda verilen örnekte olduğu gibi böyle bir durumla karşılaşılabilir.

Örnek. Sade eylemlerin uzayı $\mathcal{A} = \{a_1, a_2, a_3, \ldots\}$ sayılabilir sonsuz sayıda eylemleri içersin, $\Theta = \{\theta_1, \theta_2\}$ olsun. Kayıplar da

$$
\begin{array}{rrc}
L(\theta_1, a_1) = & L(\theta_2, a_1) = & 1 \\
L(\theta_1, a_2) = & L(\theta_2, a_2) = & \frac{1}{2} \\
L(\theta_1, a_3) = & L(\theta_2, a_3) = & \frac{1}{3} \\
\vdots & \vdots & \vdots \\
L(\theta_1, a_i) = & L(\theta_2, a_i) = & \frac{1}{i} \\
\vdots & \vdots & \vdots
\end{array}
$$

olsun. Bu eylemlerin her iki doğa durumundaki kayıp değerleri Şekil'de gösterilmiştir. Konveks kabuk $(0,0)$ noktasından $(1,1)$ noktasına ancak $(0,0)$ noktasını içermeyen doğru parçasıdır. Bu kümede $(0,0)$ noktasına istenildiğince yaklaşılabilir ama bu noktaya ulaşılamaz- böyle bir eylem yoktur- dolayısıyla bulunacak herhangi bir eyleme baskın $(0,0)$ noktasına daha yakın bir başka eylem vardır, fakat kabul edilebilir bir eylem yoktur.

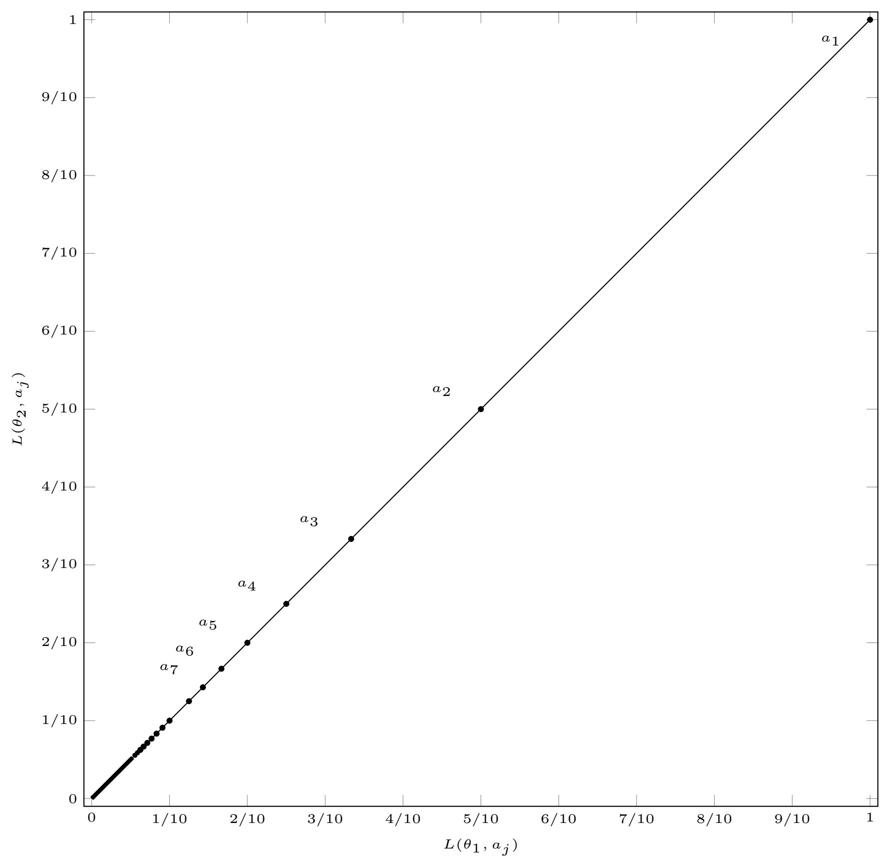
Soru. Kabul edilebilir iki eylemin herhangi bir karması olan eylem her zaman kabul edilebilir midir? İki durumun olduğu bir karar problemi için grafik üzerinde gösteriniz.

Kaynaklar

[1] D. Blackwell and M.A. Girshick (1979), Theory of games and statistical decisions, *Dover Publications*, New York.

[2] H. Chernoff and L. E. Moses (1986), Elementary Decision Theory, *Dover Publications*, New York.

[3] B. W. Lindgren (1971), Elements of Decision Theory, *Macmillan Company*, New York.



Şekil 2.5: Kabul edilebilir bir eylemin bulunmadığı bir karar problemi. Konveks kabuk üzerinde birkaç sade eylemin konumları gösterilmiştir

[4] A. Sabuncuoğlu (2004?), Analitik Geometri, 2. Baskı, *Nobel Yayınları* , Ankara?