

Sürekli Zaman Markov Karar Süreçlerinin Özgüleştirlmesi

Customizing Continuous-time Markov Decision Processes

Bora ÇEKYAY⁽¹⁾

ÖZ: Özgüleştirme tekniğinin amacı, bir üstel yarı-Markov karar sürecini (ÜYMKS) kendine özdeş, ama farklı formülasyona sahip bir başka ÜYMKS'ye dönüştürmektir. Bu sayede, özellikle optimal politikaların yapısal özelliklerini daha kolay bir şekilde ispatlamak mümkündür. Özgüleştirme tekniğinin literatürdeki mevcut hali, beklenen toplam indirgenmiş maliyeti en küçükmeye çalışan ÜYMKS'lere uygulanmaktadır. Bu makale, ÜYMKS'ler için önerilmiş olan özgüleştirme tekniğinin sürekli zaman Markov karar süreçlerine (SZMKS) nasıl uygulanabileceğini, sınırlı maliyet fonksiyonu ve üstten sınırlı geçiş hızları varsayımları altında, göstermeyi hedeflemektedir. Bu amaçla, verilen SZMKS, öncelikle bir ÜYMKS'ye dönüştürülmüştür ve daha sonra bu yeni ÜYMKS özgüleştirlmiştir.

Anahtar Kelimeler: Sürekli zaman Markov karar süreçleri, özgüleştirme, yeknesaklaştırma

Abstract: *The customization technique can convert a given exponential semi-Markov decision process (ESMDP) into another equivalent ESMDP whose formulation makes the proof of a specific structural property of the optimal policy easier. The customization technique is first proposed for ESMDPs with the expected total discounted cost criterion. This paper aims to show how the customization technique for ESMDPs can be applied to continuous-time Markov decision processes (CTMDPs) under the assumptions of bounded cost function and bounded-above transition rates. This is achieved by converting the initial CTMDP into an ESMDP which is customized later.*

Keywords: Continuous-time Markov decision processes, customization, uniformization

Jel Classification: C61, C44

1. Giriş

Markov karar süreçleri (MKS), sıralı karar verme problemlerinin çözümünde kullanılan oldukça başarılı tekniklerden birisidir. Bu problemlerde incelenen sistemin durumu, karar vericinin verdiği kararlara bağlı olarak rassal bir şekilde değişir. Bu tarz problemler, yönetim bilimleri, ekonomi ve ekoloji gibi çok farklı alanlarda ortaya çıkabilmektedir (Feinberg ve Shwartz, 2012). Karar zamanlarında sistem durumunu gözleyen karar verici, bu gözlemine bağlı olarak bir aksiyon seçer. Bu seçim neticesinde, bir maliyet (kazanç) ortaya çıkar ve sistemin bir sonraki durumu rassal olarak belirlenir. Dolayısıyla, bir MKS'yi tanımlamak için sistemin olası durumlarının, her sistem durumunda seçilebilecek aksiyonların, maliyet (kazanç) fonksiyonunun, geçiş olasılıklarının ve karar zamanları arasında geçen sürelerin tanımlanması gereklidir. Bu tanımlamalardan sonra seçilen

⁽¹⁾Doğuş Üniversitesi, Endüstri Mühendisliği Bölümü; bcekyay@dogus.edu.tr
Geliş/Received: 25-01-2018, Kabul/Accepted: 18-05-2018

çözüm sürecinde, politika iyileştirmesi (*policy improvement*), doğrusal programlama ve değer yinelemesi (*value iteration*) gibi yöntemler kullanılarak, uzun erimde toplam maliyet (kazanç) ile ilgili bir kriteri optimize eden aksiyon seçimleri belirlenir.

MKS modellerinde kullanılan iki temel yaklaşım vardır. Birinci yaklaşımda sistemin kesikli zaman anlarında gözlemlendiği varsayılır. Bu gözlem anlarına, karar zamanları, iki ardışık karar zamanı arasında geçen süreye ise karar dönemi denir. Karar dönemleri, rassal veya deterministik olabilir. Karar dönemleri önceden belirlenmiş deterministik değerler olan MKS'lere, kesikli zamanlı Markov karar süreçleri (KZMKS) denir. Karar dönemleri rassal değerler alıyorsa, bu tip MKS'lere, yarı-Markov karar süreçleri (YMKS) denir. Karar dönemlerinin üstel dağılıma uyduğu YMKS'ler, üstel yarı-Markov karar süreci (ÜYMKS) olarak isimlendirilir. MKS'lerin modellenmesinde kullanılan ikinci yaklaşımda ise sistemin sürekli gözlemlendiği ve kararların istenen herhangi bir zamanda verilebileceği varsayılır. Ayrıca, bu yaklaşımda sistem durumundaki ardışık değişimler arasındaki sürenin, üstel dağılıma uyduğu kabul edilir. Bu tarz MKS'lere, sürekli zaman Markov karar süreçleri (SZMKS) denir. (MKS'lerin detaylı incelemesi için bkz., Puterman (2005), Hu ve Yue (2007) ve Guo ve Hernández-Lerma (2009)).

Üstten sınırlı geçiş hızlarına sahip SZMKS'ler ile ÜYMKS'ler, yeknesaklaştırma (*uniformization*) olarak Türkçe'ye tercüme edilebilecek bir yöntem kullanılarak, belli varsayımlar altında, kendilerine denk olan KZMKS'lere dönüştürülebilirler. Bu sayede KZMKS'ler için geliştirilmiş olan tüm teknikler, SZMKS ve ÜYMKS'ler için de kullanılabilir. ÜYMKS'ler ile KZMKS'ler arasındaki denklik, ilk olarak Lippman (1975) tarafından fark edilmiştir. Bu denkliğin matematiksel olarak formel bir çerçeve içinde ortaya koyulması ise Serfozo (1979) tarafından yapılmıştır. Kakumanu (1977), benzer bir denkliği SZMKS'ler ile KZMKS'ler arasında tanımlamıştır. Ayrıca, genel geçiş zamanlarına sahip YMKS'ler için yeknesaklaştırma neticeleri Beutler ve Ross (1987) çalışmasında verilmiştir. Yeknesaklaştırma yönteminde orijinal MKS'nin geçiş hızları arttırılarak tek bir değere eşitlenir. Dolayısıyla dönüşüm sonucunda elde edilen MKS'nin tüm geçiş hızları aynıdır. İki MKS arasındaki denklik, maliyetlerin uygun bir şekilde değiştirilmesiyle ve bir durumdan kendine gerçekleşen yapay geçişler tanımlayarak elde edilmektedir. Yeknesaklaştırma neticesinde elde edilen yeni MKS'nin matematiksel analizi genellikle daha kolay olmaktadır. Yeknesaklaştırma tekniği özellikle monoton özelliklere sahip optimal politikaların varlığını ispatlamada oldukça faydalı olmaktadır.

Çekyay (2018), yeknesaklaştırma tekniğinde kullanılan fikri geliştirerek özgüleştirme tekniğini önermiştir. Özgüleştirme tekniği, yeknesaklaştırma tekniğinden farklı olarak geçiş hızları üstten sınırsız olduğu durumlarda da uygulanabilmektedir. Ayrıca, yeknesaklaştırma kullanıldığı halde ispatlanması kolaylaşmayan kimi yapısal neticelerin ispatını oldukça kolaylaştırabilmektedir. Çekyay ilgili çalışmasında, ÜYMKS'lerin özgülendirilmesini ilk defa, sadece durağan politikaları kullanarak ve beklenen toplam indirgenmiş maliyet amaç fonksiyonuna odaklanarak incelemiş, önerdiği özgülendirme tekniğinin SZMKS'lere de uygulanabileceğini belirtmiş ama bunun nasıl yapılacağını detaylandırmamıştır. Bu çalışmada ise, özgülendirme tekniğinin SZMKS'lere nasıl uygulanabileceği detayları ile verilecektir. Bu sayede makalenin devam eden

bölmelerinde verilecek olan varsayımları sağlayan bir ÜYMKS, geçiş hızları karar verici tarafından belirlenen bir ÜYMKS'ye dönüştürülerek analiz edilebilecektir.

Makalenin geri kalanı şu şekilde organize edilmiştir. Bölüm 2'de üzerinde çalışılacak olan ÜYMKS ve SZMKS, matematiksel olarak tanımlanacaktır. Bölüm 3'te ÜYMKS'ler için önerilmiş olan özgülleştirme metodunun SZMKS'lere nasıl uygulanabileceği gösterilecektir. Bölüm 4'te bir önceki bölümde verilmiş olan teorik neticeler, sayısal bir örneğe uygulanacaktır. Makale, Bölüm 5'te verilecek olan gelecekte yapılabilecek araştırma önerileri ve yorumlarla sonlandırılacaktır.

Makale boyunca $\mathbb{R}^{\geq 0}$, $\mathbb{R}^{> 0}$, $\mathbb{N}^{> 0}$ ve \mathbb{N} , sırasıyla negatif olmayan reel sayıları, pozitif reel sayıları, pozitif tamsayıları ve doğal sayıları göstereceklerdir.

2. Tanımlar

Bu bölümde makalenin geri kalanında kullanılacak olan MKS'ler, detaylı bir şekilde tanımlanacaktır.

Öncelikle, çalışmada kullanılacak $Y = (S, A, c, \lambda, p, \alpha)$ ÜYMKS'yi tanımlayalım. Burada S , sayılabilir durum uzayını ve A , rastgele seçilmiş aksiyon uzayını göstermektedir. Tüm maliyetler, $\alpha \in \mathbb{R}^{> 0}$ oranında sürekli olarak indirgenmektedir. Karar süreci, belli bir karar zamanında $i \in S$ durumunda gözlemlendikten sonra aksiyon uzayından bir $a \in A$ aksiyonu seçilir ve toptan $c_a(i)$ maliyeti ödenir. Karar verici, bu seçtiği aksiyonu bir sonraki karar zamanına kadar değiştiremez. Bu çalışmada $c_a(i)$ maliyetinin düzgün sınırlı olduğunu yani belli bir $M_c \in \mathbb{R}^{> 0}$ için $\sup_{i \in S, a \in A} |c_a(i)| \leq M_c$ olduğunu varsayıyoruz. Karar sürecinin i durumunda kaldığı süre, $\lambda_a(i) \in \mathbb{R}^{> 0}$ parametrelili üstel dağılıma uymaktadır ve bu sürenin sonunda karar süreci, $p_a(i; j)$ olasılığıyla $j \in S$ durumuna atlar. Bu çalışmada her $i \in E$ ve $a \in A$ için $p_a(i; i) = 0$ olduğu varsayılmıştır. Bu çalışmanın diğer bir varsayımı ise belli bir $M_\lambda \in \mathbb{R}^{> 0}$ için $\sup_{i \in S, a \in A} \lambda_a(i) \leq M_\lambda$ olmasıdır. Karar süreci Y 'nin yaptığı n . geçişin zamanı T_n ile ve Y 'nin bu geçiş sonunda bulunacağı durum Y_n ile gösterilecektir. Karar süreci Y 'nin karar dönemlerinin uzunlukları ise her $n \in \mathbb{N}^{> 0}$ için $S_n = T_n - T_{n-1}$ olarak tanımlanmıştır ve burada $T_0 = 0$ 'dir. Her bir S_n , üstel dağılıma sahiptir. Ayrıca, eğer seçilen aksiyonlar, sadece karar sürecinin karar zamanındaki durumlarına bağlıysa, iki farklı karar döneminin uzunluğu birbirinden bağımsızdır.

Şimdi de çalışmada kullanılacak olan SZMKS $Y^c = (S, A, c, q, \alpha)$ 'yi tanımlayalım. Y^c , bir SZMKS olduğu için Y 'den farklı olarak karar verici, her zaman anında seçtiği aksiyonu değiştirebilir. Eğer seçili karar değiştirilmezse karar sürecinin durumunun değiştiği iki zaman anı arasında geçen süre üstel dağılıma uymaktadır. Karar süreci i durumundayken a aksiyonu seçili ise δt süre içinde sürecin durumunun j 'ye dönüşme ihtimali yaklaşık olarak $q_a(i; j)\delta t$ 'dir. Karar sürecinin durumu i ve seçili aksiyon a iken bir sonraki durum değişim zamanına kadar geçen süre $q_a^T(i)$ parametrelili üstel dağılıma uyar ve bu süre sonunda sürecin yeni durumunun j olma olasılığı $q_a(i; j)/q_a^T(i)$ 'dir. Burada $q_a^T(i) = \sum_{j \in S} q_a(i; j)$ olarak tanımlanmıştır ve belli bir $M_q \in \mathbb{R}^{> 0}$ için $\sup_{i \in S, a \in A} q_a^T(i) \leq M_q$

olduğu varsayılmıştır. Bunun yanında bu çalışmada $q_a(i; i) = 0$ ve $q_a^T(i) \in \mathbb{R}^{>0}$ olduğu varsayılmıştır.

Markov karar süreçlerinin amacı, belirlenmiş olan amaç fonksiyonunu eniyileyen politikayı bulmaktır. SZMKS'ler için politika, her zaman anında seçilen aksiyonu belirleyen bir kuraldır. Bir SZMKS için uygulanan bir f politikası, karar sürecinin tüm geçmişine bağlı olabilir ve genel olarak, her an için bir aksiyonu seçmek yerine A kümesi üzerinde bir olasılık dağılımı tanımlar. Bu çalışmada f politikasının sadece karar sürecinin şimdiki durumuna bağlı olduğunu varsayıyoruz. Böyle bir politika $\{k_{ia}(t)\}$ şeklindeki bir fonksiyonlar kümesi ile belirlenebilir. Burada, her $i \in S, a \in A$ ve $t \in \mathbb{R}^{\geq 0}$ için $k_{ia}(t) \in \mathbb{R}^{\geq 0}$ ve $\sum_a k_{ia}(t) = 1$ olmalıdır. $k_{ia}(t)$ fonksiyonu, karar süreci i durumundayken t anında a aksiyonunun seçilme olasılığı olarak yorumlanır. Bu koşulları sağlayan bir f politikasına rassal politika denir. Eğer her $i \in S$ ve $a \in A$ için $k_{ia}(t)$ fonksiyonu t üzerinde ölçülebilir ise f politikasına ölçülebilir politika denir. Karar sürecinin sadece şimdiki durumuna bağlı olan ölçülebilir rassal politikalara ise Markov politikası denir. Bir Markov f politikasında her $i \in S, a \in A$ ve $t \in \mathbb{R}^{\geq 0}$ için $k_{ia}(t) = k_{ia}$ ise bu f politikası “durağan politika”dır ve durağan politikaların A kümesi üzerinde tanımladığı olasılık dağılımı zamandan bağımsızdır. Bir durağan politikada her $i \in S$ ve $a \in A$ için $k_{ia} = 0$ veya 1 ise bu politikaya deterministik politika denir.

Bir ÜYMKS düşünüldüğünde, herhangi bir $n \in \mathbb{N}$ için T_n karar anında seçilen aksiyon, takip eden karar dönemi boyunca T_{n+1} karar zamanına kadar değişmemektedir. Bu sebeple ÜYMKS'ler için kullanılan politikalarda zaman parametreleri doğal sayılardır. Bir ÜYMKS'de kullanılacak olan f politikası, her $n \in \mathbb{N}$ için $\{y_{ia}(n)\}$ şeklindeki fonksiyonlar kümesi ile tanımlanabilir. $y_{ia}(n)$ fonksiyonu, karar süreci i durumundayken n . karar zamanında a aksiyonunun seçilme olasılığıdır. Burada, her $i \in S, a \in A$ ve $n \in \mathbb{N}$ için $y_{ia}(n) \in \mathbb{R}^{\geq 0}$ ve $\sum_a y_{ia}(n) = 1$ olmalıdır. Eğer her $i \in S, a \in A$ ve $t \in \mathbb{R}^{\geq 0}$ için $y_{ia}(n) = y_{ia}$ ise bu f politikasına durağan politika denir. Ek olarak, eğer her $i \in S$ ve $a \in A$ için $y_{ia} = 0$ veya 1 ise bu durağan f politikasına deterministik politika denir.

Dikkat edilecek olursa ÜYMKS'ler için tanımlanabilecek tüm deterministik durağan politikalar kümesi ile SZMKS'ler için tanımlanabilecek tüm deterministik durağan politikalar kümesi birbirleriyle özdeşirler. Bu çalışmada sadece deterministik durağan politikalar üzerinde çalışılacağı için hem ÜYMKS'ye hem de SZMKS'ye ait deterministik durağan politikalar için notasyon ayırımına gidilmeyecek ve her iki politika için de f harfi kullanılacaktır. Deterministik durağan f politikasına göre herhangi bir $i \in E$ durumunda seçilen aksiyon, $f(i)$ ile gösterilecektir. Burada, her $i \in E$ için $f(i) \in A$ olduğu açıktır.

Bu çalışmada özgülleştirme tekniği, Çekyay, (2018)'de yapıldığı gibi, deterministik durağan politikalar ve beklenen toplam indirgenmiş maliyet kriteri için sunulacaktır. İncelediğimiz ÜYMKS, Y , belli bir deterministik durağan f politikasına göre yönetildiğinde ve başlangıç durumu i olduğunda, beklenen toplam indirgenmiş maliyet fonksiyonunu

$$W_{f,\alpha}(i) = \lim_{N \rightarrow \infty} E_f^i \left[\sum_{n=0}^N e^{-\alpha T_n} c_{f(Y_n)}(Y_n) \right] = E_f^i \left[\sum_{n=0}^{\infty} e^{-\alpha T_n} c_{f(Y_n)}(Y_n) \right] \quad (1)$$

şeklinde tanımlıyoruz. Bu ifadede beklenen değer sembolünün üst indisi karar sürecinin başlangıç durumunu, alt indisi ise uygulanan politikayı göstermektedir. Aynı deterministik durağan f politikası altında incelenen SZMKS'nin, Y^c 'nin, beklenen toplam indirgenmiş maliyet fonksiyonunu ise

$$W_{f,\alpha}^c(i) = E_f \left[\int_0^\infty e^{-\alpha t} c_{f(Y_t^c)}(Y_t^c) \mid Y_0^c = i \right] \quad (2)$$

şeklinde tanımlıyoruz. Burada Y_t^c, Y^c karar sürecinin $t \in \mathbb{R}^{\geq 0}$ anındaki durumudur.

Teorem 1, Denklem (1)'de tanımlanan maliyet fonksiyonunu rahatlıkla kullanabileceğimizi ve Çekyay (2018)'deki 2.1 numaralı temel varsayımı sağladığımızı göstermektedir. Bu neticenin ispatlanabilmesi için ilk önce aşağıdaki basit önsavın ispatlanması gereklidir.

Önsav 1 Her $n \in \mathbb{N}$ için $E_f^i[e^{-\alpha S_{n+1}}] \leq \frac{M_\lambda}{\alpha + M_\lambda}$.

İspat: Eğer $Y_n = j$ ise $S_{n+1}, \lambda_{f(j)}(j)$ parametrelü üstel dağılıma uyar ve dolayısıyla, $E_f^j[e^{-\alpha S_{n+1}} \mid Y_n = j], \lambda_{f(j)}(j)$ parametrelü üstel dağılımın Laplace dönüşümü olur. Bu durumda,

$$\begin{aligned} E_f^i[e^{-\alpha S_{n+1}}] &= \sum_{j \in \mathcal{S}} E_f^j[e^{-\alpha S_{n+1}} \mid Y_n = j] P_f^i\{Y_n = j\} \\ &= \sum_{j \in \mathcal{S}} \frac{\lambda_{f(j)}(j)}{\alpha + \lambda_{f(j)}(j)} P_f^i\{Y_n = j\} \\ &\leq \sum_{j \in \mathcal{S}} \frac{M_\lambda}{\alpha + M_\lambda} P_f^i\{Y_n = j\} = \frac{M_\lambda}{\alpha + M_\lambda} \end{aligned}$$

olur.

Önsav 2. Her $n \in \mathbb{N}$ için $E_f^i[e^{-\alpha T_n}] \leq \left(\frac{M_\lambda}{\alpha + M_\lambda}\right)^n$.

İspat: $T_0 = 0$ olduğu için $n = 0$ olduğunda önsavdaki ifadenin doğru olduğu açıktır. Eğer $n \geq 1$ ise $T_n = S_1 + \dots + S_n$ olur ve bu toplamda S_i 'ler birbirlerinden bağımsızdırlar. Dolayısıyla, T_n 'in Laplace dönüşümü, S_i 'lerin Laplace dönüşümlerinin çarpımına eşit olur. Bu durumda:

$$E_f^i[e^{-\alpha T_n}] = E_f^i[e^{-\alpha S_1}] E_f^i[e^{-\alpha S_2}] \dots E_f^i[e^{-\alpha S_n}] \leq \left(\frac{M_\lambda}{\alpha + M_\lambda}\right)^n$$

olur. Bu çıkarımdaki eşitsizlik, Önsav 1'in sonucudur.

Teorem 1. Denklem (1)'de tanımlanan maliyet fonksiyonu $W_{f,\alpha}(i)$, iyi tanımlı bir fonksiyondur, her zaman sonlu değerler alır ve Çekyay (2018)'deki 2.1 numaralı temel varsayımı sağlar.

İspat: Öncelikle (1)'deki limitin var ve sonlu olduğunu göstereceğiz. Bunun için $\sum_{n=0}^\infty E_f^i[e^{-\alpha T_n} c_{f(Y_n)}(Y_n)]$ 'nin mutlak yakınsak bir sonsuz dizi olduğunu göstermemiz yeterlidir. Mutlak değer, konveks bir fonksiyon olduğu için Jensen'nin eşitsizliğine göre $|E_f^i[e^{-\alpha T_n} c_{f(Y_n)}(Y_n)]| \leq E_f^i[e^{-\alpha T_n} |c_{f(Y_n)}(Y_n)|] \leq M_c E_f^i[e^{-\alpha T_n}]$ olur. Dolayısıyla,

$$\begin{aligned} \sum_{n=0}^{\infty} |E_f^i[e^{-\alpha T_n} c_{f(Y_n)}(Y_n)]| &\leq M_c \sum_{n=0}^{\infty} E_f^i[e^{-\alpha T_n}] \leq M_c \sum_{n=0}^{\infty} \left(\frac{M_\lambda}{\alpha + M_\lambda}\right)^n \\ &= \frac{M_c(\alpha + M_\lambda)}{\alpha} < \infty \end{aligned} \quad (3)$$

olur. Bu çıkarımdaki ikinci eşitsizlik, Önsav 2'den gelmektedir. Mutlak yakınsak sonsuz bir dizi, yakınsak olduğu için Denklem (1)'deki limit vardır ve her zaman sonludur. Bunun sonucu olarak $W_{f,\alpha}(i)$, iyi tanımlı bir fonksiyondur. $c_a^+(i) = \max\{c_a(i), 0\}$ ve $c_a^-(i) = \max\{-c_a(i), 0\}$ tanımlamalarını yapalım. Açık ki $\sup_{i \in S, a \in A} |c_a^+(i)| \leq M_c$ ve $\sup_{i \in S, a \in A} |c_a^-(i)| \leq M_c$ olur. Dolayısıyla, bu ispatın yukarıdaki adımlarını aynen takip ederek $\sum_{n=0}^{\infty} E_f^i[e^{-\alpha T_n} c_{f(Y_n)}^+(Y_n)] < \infty$ ve $\sum_{n=0}^{\infty} E_f^i[e^{-\alpha T_n} c_{f(Y_n)}^-(Y_n)] < \infty$ olduğu gösterilebilir. Bunun sonucu olarak (Çekyay 2018)'deki 2.1 numaralı temel varsayım, $W_{f,\alpha}(i)$ için sağlanmış olur.

Teorem 1 sayesinde $W_{f,\alpha}(i)$ fonksiyonun, iyi tanımlı bir fonksiyon olduğunu ve üzerinde rahatlıkla çalışabileceğimizi görmüş olduk. Ayrıca, aynı şekilde, Çekyay (2018)'deki tüm neticelerin bu çalışma kapsamında kullanılabilmesi delillendirilmiştir.

3. SZMKS'lerin ÜYMKS'ye Dönüştürülerek Özgüleştirmesi

Bu bölümde Çekyay (2018) tarafından önerilmiş olan özgüleştirme tekniğinin SZMKS'ler için nasıl uygulanabileceği detaylarıyla gösterilecektir. Bunun için ilk önce Bölüm 2'de verilmiş olan tanıma uyan bir SZMKS, bir ÜYMKS'ye dönüştürülecektir. Daha sonra Çekyay (2018) tarafından önerilen özgüleştirme tekniği, elde edilmiş olan ÜYMKS'ye uygulanacaktır.

Varsayalım ki elimizde Bölüm 2'de tanımlandığı gibi bir SZMKS $Y^c = (S, A, c, q, \alpha)$ olsun. Bu karar sürecinin parametrelerini kullanarak yeni bir ÜYMKS $\bar{Y} = (S, A, \bar{c}, \bar{q}, \bar{p}, \alpha)$ tanımlayalım. Bu tanımlamada $\bar{q}_a(i) = q_a^T(i)$ ve

$$\bar{c}_a(i) = \frac{c_a(i)}{\alpha + q_a^T(i)}, \quad \bar{p}_a(i; j) = \begin{cases} 0, & i = j \\ q_a(i; j), & i \neq j. \end{cases} \quad (4)$$

Bu yeni tanımlanan ÜYMKS'ye, \bar{Y} 'ye, ait olan ve (1)'e göre tanımlanan beklenen toplam indirgenmiş maliyet fonksiyonunu $\bar{W}_{f,\alpha}(i)$ ile gösterelim. (Feinberg, 2004) çalışmasındaki Teorem 4.5, yukarıda tanımlanmış iki karar sürecinin deterministik durağan politikalar altında birbirlerine denk olduklarını göstermektedir. Diğer bir ifade ile herhangi bir deterministik durağan f politikası için $\bar{W}_{f,\alpha}(i) = W_{f,\alpha}^c(i)$.

Bu aşamada Çekyay (2018)'de önerildiği şekliyle \bar{Y} sürecini özgüleştirebiliriz. Bunun için

$\tilde{\lambda}_a(i) \geq \bar{\lambda}_a(i)$ koşulunu sağlayacak şekilde yeni $\tilde{\lambda}$ geçiş hızlarının seçildiğini varsayalım. Bu yeni geçiş hızlarını kullanarak yeni geçiş olasılıklarını ve maliyet fonksiyonunu aşağıdaki gibi hesaplayalım:

$$\tilde{c}_a(i) = \bar{c}_a(i) \frac{\alpha + \bar{\lambda}_a(i)}{\alpha + \tilde{\lambda}_a(i)}, \quad \tilde{p}_a(i; j) = \begin{cases} 1 - \frac{\bar{\lambda}_a(i)}{\tilde{\lambda}_a(i)}, & i = j \\ \frac{\bar{\lambda}_a(i) \bar{p}_a(i; j)}{\tilde{\lambda}_a(i)}, & i \neq j. \end{cases} \quad (5)$$

Böylece $\tilde{Y} = (S, A, \tilde{c}, \tilde{\lambda}, \tilde{p}, \alpha)$ ÜYMKS'si tanımlanmış oldu. Bu yeni karar sürecinin (1)'e göre tanımlanan beklenen toplam indirgenmiş maliyet fonksiyonunu $\tilde{W}_{f,\alpha}(i)$ ile gösterelim. Çekyay (2018)'de ispatlanan Teorem 2.3, herhangi bir deterministik durağan f politikası için $\tilde{W}_{f,\alpha}(i) = \bar{W}_{f,\alpha}(i)$ olduğunu göstermektedir.

Bu noktaya kadar yapılan analiz neticesinde aşağıdaki teorem ispatlanmış oldu.

Teorem 2. $Y^c = (S, A, c, q, \alpha)$, Bölüm 67'deki varsayımları sağlayan bir SZMKS olsun. $\tilde{Y} = (S, A, \tilde{c}, \tilde{\lambda}, \tilde{p}, \alpha)$ ise geçiş hızları $\infty > \tilde{\lambda}_a(i) \geq q_a^T(i)$ varsayımı altında serbestçe seçilmiş ve diğer parametreleri, her $i, j \in S$ ve $a \in A$ için

$$\tilde{c}_a(i) = \frac{c_a(i)}{\alpha + \tilde{\lambda}_a(i)}, \quad \tilde{p}_a(i; j) = \begin{cases} 1 - \frac{q_a^T(i)}{\tilde{\lambda}_a(i)}, & i = j \\ \frac{q_a(i; j)}{\tilde{\lambda}_a(i)}, & i \neq j. \end{cases} \quad (6)$$

eşitlikleri ile hesaplanan bir ÜYMKS olsun. Eğer Y^c ve \tilde{Y} , aynı deterministik durağan f politikası ile kontrol ediliyorlarsa, her $i \in S$ için $W_{f,\alpha}^c(i) = \tilde{W}_{f,\alpha}(i)$.

İspat: Bu bölümün başından beri yapılan açıklamalara ek olarak sadece Denklem 6'nın, Denklem 4 ve Denklem 5'ten çıkarılabileceğinin gösterilmesi gereklidir. Bunun için de $\tilde{\lambda}_a(i) = q_a^T(i)$ eşitliğini kullanmak yeterlidir.

Teorem 2'nin doğrudan bir sonucu aşağıda verilmiştir.

Sonuç 1. Eğer hem Y^c hem de \tilde{Y} için beklenen toplam indirgenmiş maliyet fonksiyonunu deterministik durağan politikalar eniyiliyorsa, her iki karar süreci için de aynı deterministik durağan politika optimaldir.

SZMKS'ler ve ÜYMKS'ler için deterministik durağan politikalar, her zaman optimal olmak zorunda değildir. SZMKS'ler için Guo ve Hernández-Lerma (2009) (Bölüm 4.5 ve 4.6), ÜYMKS'ler için Puterman (2005) (Bölüm 6.2.4), deterministik durağan politikaların optimalliğini garanti eden koşullardan bazılarını vermektedirler. Bu koşullar altında Sonuç 1, elimizdeki SZMKS ile bu süreci özgülleştirerek elde ettiğimiz ÜYMKS'yi aynı deterministik durağan politikaların optimize edeceğini söylemektedir. Örneğin, S ve A

sonlu kümeler olduklarında, hem Y^c hem de \tilde{Y} için aynı deterministik durağan politika optimaldir.

4. Sayısal Örnek

Bu bölümde bir önceki bölümde ispatlanmış olan neticenin sayısal bir örnek üzerinde uygulaması yapılacaktır. Bunun için basit bir $Y^c = (S, A, c, q, \alpha)$ SZMKS tanımlayalım. Bu karar süreci için $S = \{1,2,3,4\}$, $A = \{0,1\}$, $\alpha = 0,1$, $c_0 = [4 \ 5 \ 6 \ 7]$, $c_1 = [1 \ 3 \ 8 \ 9]$,

$$q_0 = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 2 & 1 & 0 & 1 \\ 2 & 2 & 2 & 0 \end{bmatrix} \text{ ve } q_1 = \begin{bmatrix} 0 & 2 & 2 & 2 \\ 2 & 0 & 2 & 1 \\ 1 & 2 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

olsun. Bu durumda $q_0^T = [2 \ 3 \ 4 \ 6]$ ve $q_1^T = [6 \ 5 \ 4 \ 3]$ olur.

Varsayalım ki optimal politikanın bazı özelliklerini ispatlamak için q_0^T ve q_1^T vektörlerinin değerlerinin sırasıyla $[3 \ 3 \ 6 \ 6]$ ve $[6 \ 6 \ 4 \ 4]$ şeklinde olmaları işimizi kolaylaştıracak. Bu durumda tanımladığımız SZMKS'yi özgüleştirerek elde edeceğimiz yeni $\tilde{Y} = (S, A, \tilde{c}, \tilde{\lambda}, \tilde{p}, \alpha)$ ÜYMKS'nin geçiş hızlarının $\tilde{\lambda}_0 = [3 \ 3 \ 6 \ 6]$ ve $\tilde{\lambda}_1 = [6 \ 6 \ 4 \ 4]$ şeklinde olması gerekir. Dikkat edilecek olursa bu geçiş hızları $\tilde{\lambda} \geq q^T$ olacak şekilde seçilmiştir. Bu durumda Denklem 6'ya göre $\tilde{c}_0 = [1,29 \ 1,61 \ 0,98 \ 1,15]$, $\tilde{c}_1 = [0,16 \ 0,49 \ 1,95 \ 2,20]$,

$$\tilde{p}_0 = \begin{bmatrix} 1/3 & 1/3 & 1/3 & 0 \\ 1/3 & 0 & 1/3 & 1/3 \\ 1/3 & 1/6 & 1/3 & 1/6 \\ 1/3 & 1/3 & 1/3 & 0 \end{bmatrix} \text{ ve } \tilde{p}_1 = \begin{bmatrix} 0 & 1/3 & 1/3 & 1/3 \\ 1/3 & 1/6 & 1/3 & 1/6 \\ 1/4 & 1/2 & 0 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{bmatrix}$$

olur. Sırasıyla, (Guo ve Hernández-Lerma 2009)'de yer alan Teorem 4.10 ve (Puterman 2005)'de yer alan Teorem 11.3.2'ye göre hem Y^c için hem de \tilde{Y} için optimal olan deterministik durağan politikalar vardır. Dolayısıyla, **Teorem 2** ve **Sonuç 1**'e göre aynı deterministik durağan f politikası, iki karar süreci için de optimal olmalıdır ve bu politika için $W_{f,\alpha}^c(i) = \tilde{W}_{f,\alpha}(i)$ olmalıdır.

Bu sonucu doğrulamak için her iki karar sürecinin optimal politikalarını ayrı ayrı sayısal olarak bulacağız. SZMKS Y^c 'nin optimal politikasını bulmak için (Guo ve Hernández-Lerma, 2009)'de yer alan Teorem 4.14 verilmiş olan değer yinelemesi algoritması kullanılmıştır. Optimal politika $f = [1 \ 1 \ 0 \ 0]$ ve bu politikaya karşılık gelen beklenen toplam indirgenmiş maliyet vektörü $W_{f,\alpha}^c = [41,75 \ 41,92 \ 42,41 \ 42,49]$ olarak bulunmuştur. ÜYMKS \tilde{Y} 'nin optimal politikasını bulmak için Puterman (2005)'de içindeki Bölüm 11.3.4'de bahsedilen değer yinelemesi algoritması kullanılmıştır. Optimal politika $f = [1 \ 1 \ 0 \ 0]$ ve bu politikaya karşılık gelen beklenen toplam indirgenmiş maliyet vektörü $\tilde{W}_{f,\alpha} = [41,75 \ 41,92 \ 42,41 \ 42,49]$ olarak bulunmuştur. Her iki karar sürecinin, f politikası altında, beklenen toplam indirgenmiş maliyet kriteri açısından, birbirine denk olduğu aşikardır.

5. Sonuç

MKS'ler sıralı karar verme konusunda sıklıkla kullanılan araçlardan birisidir. MKS'ler kullanılarak incelenen problemin sayısal çözümü yapılabileceği gibi optimal politikanın belli yapısal özelliklere sahip olduğu ispatlanabilir. Bu matematiksel ispatları kolaylaştırmak adına Lippman (1975) çalışmasından bu yana yeknesaklaştırma tekniği çoğu kez kullanılan önemli bir yardımcıdır. Bu teknikte orijinal MKS'nin geçiş hızları sabit bir değere dönüştürülmektedir. Bu dönüşüm daha basit bir formülasyon ile çalışmaya olanak tanımaktadır ama son durumdaki geçiş hızlarının hepsinin aynı olmasını ve orijinal MKS'nin geçiş hızlarının üstten sınırlı olmasını şart koşmaktadır. Çekyay (2018) tarafından önerilen özgülleştirme tekniğinde ise son durumdaki geçiş hızları farkı değerler alabilmektedir ve orijinal geçiş hızları, üstten sınırsız olabilmektedir. Bu özellikleri sayesinde yeknesaklaştırmanın yardımcı olamadığı durumlarda faydalı olabilmektedir. Özgülleştirme tekniği ilk olarak ÜYMKS'ler için tanımlanmıştır. Bu çalışma, özgülleştirme tekniğinin SZMKS'lere nasıl uygulanabileceğini göstermektedir. Bunun için verilen SZMKS, öncelikle bir ÜYMKS'ye çevirilmiştir ve bu yeni sürece özgülleştirme tekniği uygulanmıştır.

Bu çalışma ve Çekyay (2018), sadece beklenen toplam indirgenmiş maliyet amaç fonksiyonuna odaklanmıştır. Ortalama maliyet gibi diğer amaç fonksiyonları için özgülleştirme neticelerinin çıkarılması, ileriki çalışmalara bırakılmıştır. Bunun yanında, SZMKS kullanılarak modellenmiş bir problem için optimal politikanın yapısının araştırılmasında, özgülleştirmenin ne şekilde faydalı olabileceğinin gösterimi de gelecekte çalışılacak projeler arasına eklenmiştir.

6. Kaynakça

- Beutler, F. J., ve Keith, W. R. (1987). Uniformization for semi-Markov decision processes under stationary policies. *Journal of Applied Probability* 24 (3). Cambridge University Press, 644–56.
- Çekyay, B. (2018). Customizing exponential semi-Markov decision processes under the discounted cost criterion. *European Journal of Operational Research* 266 (1), 168–78. doi:10.1016/j.ejor.2017.09.016.
- Feinberg, E. A., Shwartz, A., ed. (2012). *Handbook of Markov decision processes: methods and applications (Vol. 40)*. Springer Science & Business Media.
- Feinberg, E. A. (2004). Continuous time discounted jump Markov decision processes: a discrete-event approach. *Mathematics of Operations Research* 29 (3). INFORMS, 492–524.
- Guo, X. ve Hernández-Lerma, O.. (2009). *Continuous-time Markov decision processes: theory and applications*. Stochastic Modelling and Applied Probability. Springer Verlag.
- Hu, Q., ve Yue, W. (2007). *Markov decision processes with their applications*. Vol. 14. Springer.

- Kakumanu, P. (1977). Relation between continuous and discrete time Markovian decision problems. *Naval Research Logistics Quarterly* 24 (3). Wiley Online Library, 431–39.
- Lippman, S. A., (1975). Applying a new device in the optimization of exponential queuing systems. *Operations Research* 23 (4). INFORMS, 687–710.
- Puterman, M. L., (2005). *Markov decision processes: discrete stochastic dynamic programming*. Wiley Series in Probability and Statistics. New Jersey: John Wiley and Sons.
- Serfozo, R. F., (1979). “An equivalence between continuous and discrete time Markov decision processes.” *Operations Research* 27 (3). INFORMS, 616–20.