

Klassificatie van Markov beslissingsketens

Begeleider: Lodewijk Kallenberg

Bij Markov beslissingsketens hebben we niet te maken met één Markov keten, maar met een aantal Markov ketens. We zullen veronderstellen dat de ketens eindig veel toestanden hebben. Formeel is het model als volgt:

- S is een eindige *toestandsruimte*, zeg $S = \{1, 2, \dots, N\}$;
- in toestand $i \in S$ wordt een actie uit een eindige *actieverzameling* $A(i)$ gekozen;
- indien in toestand i actie $a \in A(i)$ wordt gekozen, dan gebeurt het volgende: er is een *directe opbrengst* $r_i(a)$ en met *overgangskans* $p_{ij}(a)$ is het systeem op het volgende beslissingstijdstip in toestand j .

Een (*deterministische*) *strategie* f is een beslisregel: het geeft aan welke actie in toestand i wordt gekozen; de actie die in toestand i wordt gekozen noteren we met $f(i)$. Iedere strategie genereert een Markov keten met overgangsmatrix $P(f)$, waarbij $[P(f)]_{ij} = p_{ij}(f(i))$ voor alle $i, j \in S$.

Er zijn verschillende manieren om Markov beslissingsketens te klassificeren. Er kan onderscheid worden gemaakt tussen *communicerende* en *niet-communicerende* Markov beslissingsketens. Een Markov beslissingsketen is communicerend als voor iedere $i, j \in S$ er een strategie f bestaat (die van i en j mag afhangen) zodanig dat in de Markov keten $P(f)$ toestand j bereikbaar is vanuit toestand i .

Een Markov beslissingsketen heet *zwak communicerend* als $S = S_1 \cup S_2$ met $S_1 \cap S_2 = \emptyset$, met S_1 een gesloten communicerende verz. onder een bepaalde strategie en S_2 een (mogelijk lege) verz. toestanden die transiënt zijn onder alle strategieën.

Een tweede klassificatie betreft de ergodische structuur. We maken onderscheid tussen *irreducibele*, *unichain* en *multichain* Markov beslissingsketens. Een Markov beslissingsketen is irreducibel als de Markov keten $P(f)$ irreducibel is voor iedere strategie f ; unichain als voor iedere strategie f the Markov keten $P(f)$ precies één recurrente klasse heeft plus een (mogelijk lege) verz. van transiënte toestanden; multichain als er een strategie f bestaat waarvoor de Markov keten $P(f)$ ten minste twee recurrente klassen heeft.

De probleemstellingen van dit project, bij gegeven Markov beslissingsketen, luiden:

1. a. hoe bepaal je of het model communicerend of een niet-communicerend.
b. Als het niet-communicerend is, hoe bepaal je of het zwak communicerend is?
c. Bestaan er polynomiale algoritmen voor deze klassificatieproblemen?
d. Als je geen polynomiale algoritmen kunt vinden, zijn deze probleem dan $\mathcal{NP}$ -volledig?
2. a. Hoe bepaal je of het model een irreducibel, unichain of multichain is?
b. Bestaan er polynomiale algoritmen voor deze klassificatie?
c. Als je geen polynomiale algoritmen kunt vinden, zijn deze probleem dan $\mathcal{NP}$ -volledig?