# ΚΕΦΑΛΑΙΟ 2

## Υπολογισμός της άριστης συνάρτησης τιμών για πεπερασμένο χρονικό ορίζοντα σε μία POMDP με την μέθοδο των Smallwood –Sondik- Lovejoy.

### Περίληψη

Στο κεφάλαιο αυτό θα περιγράψουμε τον υπολογισμό της άριστης συνάρτησης τιμών $V_t$ για τον χρονικό ορίζοντα t , όταν είναι γνωστή η άριστη συνάρτηση τιμών $V_{t-1}$ για τον χρονικό ορίζοντα t-1 σε μία POMDP για πρόβλημα εσόδων. Οι Smallwood-Sondik [119] έδωσαν αρχικά μια μέθοδο υπολογισμού που στη συνέχεια τροποποιήθηκε από τον Lovejoy [77] με την διόρθωση κάποιων ατελειών της αρχικής μεθόδου. Θα αναφερόμαστε στην τελευταία ως μέθοδο των Smallwood –Sondik-Lovejoy.

Στην ενότητα 2.1 δείχνουμε ότι το πρόβλημα υπολογισμού της $V_t$ από την $V_{t-1}$ ανάγεται στον υπολογισμό της Hυ αν η υ είναι γνωστή κατά τμήματα γραμμική και κυρτή συνάρτηση (p.w.l.c). Πιο συγκεκριμένα, δείχνουμε ότι ο υπολογισμός της Hυ ανάγεται στην εύρεση του συνόλου των λειτουργικών gradient vectors της Hυ αν γνωρίζουμε το σύνολο των λειτουργικών gradient vectors της υ.

Στην ενότητα 2.2 παρουσιάζουμε τον αλγόριθμο του ενός βήματος (one-pass algorithm) των Smallwood-Sondik [117] με τον οποίο υπολογίζεται η τιμή Hυ(π) για δοσμένο δ.π. π.

Στην ενότητα 2.3 περιγράφουμε μία μέθοδο καταγραφής του συνόλου G των " εν δυνάμει" gradient vectors της συνάρτησης Hυ και στη συνέχεια μέθοδο συρρίκνωσης του G σε μια ελάχιστη αντιπροσώπευση από λειτουργικά gradient vectors της Hυ, μέσω γραμμικών προγραμμάτων.

Τέλος στην ενότητα 2.4 προσδιορίζουμε την περιοχή του χώρου Π στην οποία ένα gradient vector για την Hυ είναι λειτουργικό με την μέθοδο των Smallwood-Sondik-Lovejoy μέσω γραμμικών προγραμμάτων.

## 2.1. Εισαγωγή

Είναι γνωστό ότι η βέλτιστη συνάρτηση τιμών για το πρόβλημα εσόδων σε πεπερασμένο χρονικό ορίζοντα είναι κατά τμήματα γραμμική και κυρτή (βλέπε και ενότητα 1.5 ).Επομένως η βέλτιστη συνάρτηση τιμών για ένα πρόβλημα t-1 ορίζοντα μπορεί να εκφρασθεί :

$$V_{t-1}(\pi)=\max\ \{\pi.\gamma:\ \gamma\in\Gamma_{t-1}\},\pi\in\Pi$$

όπου $\Gamma_{t-1}=\{\gamma^1_{t-1},\gamma^2_{t-1},\ldots\gamma^k_{t-1}\}$ένα πεπερασμένο σύνολο από διανύσματα του $R^N$ που ονομάζονται ´´gradient-vectors´´ ή απλά ´´gradients´´.Τα διανύσματα $\pi\in\Pi$ θεωρούνται διανύσματα γραμμές, ενώ τα ´´gradients´´ διανύσματα στήλες.Ο αλγόριθμος των Smallwood-Sondik [119],ουσιαστικά βασίζεται στην εύρεση του $V_t(\pi)$, όταν είναι γνωστό το $V_{t-1}(\pi)$. Δηλαδή :

$$V_t(\pi)=HV_{t-1}(\pi), \tag{2.1.1}$$

όπου H είναι ο τελεστής μεγιστοποίησης (βλ. ενότητα 1.4)

Πρόκειται δηλαδή για μια επέκταση τής μεθόδου των διαδοχικών προσεγγίσεων (value-iteration) από τις Μαρκοβιανές διαδικασίες αποφάσεων (MDP) (βλέπε παράγραφο 1.1) στις μερικά παρατηρήσιμες Μαρκοβιανές διαδικασίες αποφάσεων (POMDP). Η δυσκολία, που αντιμετωπίζουμε οφείλεται κύρια στο γεγονός ότι σε μια POMDP ο χώρος των δ.π ,Π, είναι συνεχής.

Το πρόβλημα ανάγεται στον υπολογισμό της συνάρτησης Hυ αν η συνάρτηση υ είναι κατά τμήματα γραμμική και κυρτή (p.w.l.c). Τότε η υ αντιπροσωπεύεται από ένα πεπερασμένο σύνολο από ´´gradient-vectors´´,έστω Γ,

$$\Gamma=\{\gamma^1,\gamma^2,\ldots\gamma^k\}$$

$$\upsilon(\pi) = \max_{1 \le j \le k} \pi.\gamma^j , \pi \in \Pi .$$

**Ορισμός 2.1.1:** Έστω $\gamma \in \Gamma$ .Η περιοχή

R(γ,Γ)={$\pi \in \Pi$:π.γ$\ge$π.$\bar{\gamma}$, $\forall$ $\bar{\gamma} \in \Gamma$} $\neq \emptyset$,λέγεται συσχετιζόμενη με το γ περιοχή, (ή υποστηρίζουσα το gradient $\gamma$ ).

Σημειώνουμε ότι η περιοχή R(γ,Γ) είναι κυρτό πολύεδρο του Π.

Από τον ορισμό συνάγεται ότι :

Αν $\pi \in$ R(γ, Γ) τότε $\upsilon(\pi) = \pi.\gamma$ .

Προφανώς ισχύουν :

$$\bigcup_{\gamma \in \Gamma} R(\gamma,\Gamma)=\Pi \quad \text{και} \quad \text{int}(R(\gamma,\Gamma)) \bigcap \text{int}(R(\bar{\gamma},\Gamma))=\emptyset , \forall \ \gamma \neq \bar{\gamma} . \qquad \underline{\mathbf{2.1.2}}$$

Έτσι ο χώρος Π διαμερίζεται σε συσχετιζόμενες περιοχές των ΄΄gradient vectors΄΄ $\gamma \in \Gamma$ . Σημειώνουμε ακόμη ότι η συνάρτηση:

$$H\upsilon(\pi) = \max_{a \in A}\{\pi.q^a + \beta.\sum_{\theta}\{\theta / \pi, \alpha\}.\upsilon(\mathrm{T}(\pi, \theta, \alpha)\}, \pi \in \Pi, \qquad \underline{\mathbf{2.1.3}}$$

*αποδεικνύεται ότι είναι p.w.l.c, βλέπε Smallwood–Sondik [117], και επομένως αντιπροσωπεύεται από ένα πεπερασμένο σύνολο ΄΄gradient vectors΄΄ έστω* $\Gamma_H$ *,έτσι* ώστε:

$$H\upsilon(\pi) = \max \{\pi.\gamma' : \gamma' \in \Gamma_H\}. \qquad \underline{\mathbf{2.1.4}}$$

Θα μας απασχολήσουν κύρια τρία προβλήματα .

**1)**Ο υπολογισμός του Η$\upsilon$(π) για δοσμένο δ.π $\pi \in \Pi$ (ενότητα 2.2).

**2)**Ο καθορισμός του συνόλου ,$\Gamma_H$, ως ελάχιστη αντιπροσώπευση από λειτουργικά ΄΄gradient vectors΄΄ για την συνάρτηση Η$\upsilon$(π),$\pi \in \Pi$,(ενότητα 2.3).

**3)**Ο καθορισμός των συσχετιζόμενων περιοχών των λειτουργικών ΄΄gradient vectors΄΄ για την συνάρτηση Ηυ (ενότητα 2.4).

Η σχέση (2.1.3) γράφεται:

$$H\upsilon(\pi) = \max_{a \in A}\{\pi.q^a + \beta.\sum_{\theta}\{\theta / \pi, \alpha\}.\mathrm{T}(\pi, \theta, \alpha).\gamma^{l(\pi,\theta,\alpha)}\}$$

$$= \max_{a \in A} \{\pi.(q^\alpha + \beta.P^\alpha. \sum_{\theta} R_\theta^\alpha \, . \, \gamma^{l\,(\pi,\theta,\alpha)} )\}$$

$$= \max_{a \in A} \{\pi.\gamma_{\alpha}(\pi)\}\ ,\ \pi \in \Pi\ ,$$

όπου $$\gamma_{\alpha}(\pi) = q^{\alpha} + \beta.P^{\alpha}.\sum_{\theta} R_{\theta}^{\alpha}\ .\ \gamma^{l\,(\pi,\theta,\alpha)}\ ,\ \forall\,\pi \in \Pi,\ \forall\,\alpha \in A.$$

Ο δείκτης $l(\pi,\theta,\alpha)$ είναι ο δείκτης μεγιστοποίησης στη σχέση:

$$T(\pi,\theta,\alpha).\gamma^{l(\pi,\theta,\alpha)} = \max_{1 \le j \le k} T\,(\pi,\theta,\alpha).\gamma^{j},$$

η ισοδύναμα

$$\pi.\ P^{\alpha}.R_{\theta}^{\alpha}\ .\ \gamma^{l(\pi,\theta,\alpha)} = \max_{1 \le j \le k} \pi.\ P^{\alpha}.R_{\theta}^{\alpha}\ .\ \gamma^{j} \qquad \underline{\mathbf{2.1.5}}$$

Με τον παρακάτω αλγόριθμο $A_3$, που είναι γνωστός σαν αλγόριθμος του ενός βήματος και οφείλεται στον Sondik [119] μπορούμε να υπολογίσουμε την τιμή $H\upsilon(\tilde{\pi})$ για δοσμένο $\tilde{\pi} \in \Pi$.

## 2.2. Αλγόριθμος τού ενός βήματος

Σκοπός του αλγορίθμου του ενός βήματος (one–pass algorithm) είναι η εύρεση λειτουργικού «gradient» της Ηυ σε κάποιο δ.π , $\tilde{\pi}$, καθώς και της βέλτιστης απόφασης στο συγκεκριμένο $\tilde{\pi}$.Η ονομασία «αλγόριθμος του ενός βήματος», σχετίζεται με την εφαρμογή του στο ενός βήματος πέρασμα από τον χρονικό ορίζοντα t-1 στον χρονικό ορίζοντα t.

### Αλγόριθμος ( τού ενός βήματος ) $A_3$ ( Sondik [119] )

**ΒΗΜΑ 1:** Για $\theta \in \Theta$, $\alpha \in A$ βρίσκουμε τον δείκτη $l(\tilde{\pi},\theta,\alpha)$ έτσι ώστε :

$$\tilde{\pi}.P^{a}.R_{\theta}^{\alpha}.\gamma^{l(\tilde{\pi},\theta,\alpha)} = \max_{1 \le j \le k} \tilde{\pi}.P^{\alpha}.R_{\theta}^{\alpha}\ \gamma^{j}.$$

**ΒΗΜΑ 2:** Υπολογίζουμε τα gradients :

$$\gamma_{\alpha}(\tilde{\pi}) = q^{\alpha} + \beta.P^{\alpha}.\sum_{\theta} R_{\theta}^{\alpha}\ \gamma^{l(\tilde{\pi},\alpha,\theta)}\ ,\alpha \in A.$$

**ΒΗΜΑ3:** Υπολογίζουμε την ποσότητα $H\upsilon(\tilde{\pi})$ από τη σχέση:

$$H\upsilon(\tilde{\pi}) = \max_{\alpha} \tilde{\pi}.\gamma_{\alpha}(\tilde{\pi}) = \tilde{\pi}.\gamma_{\alpha^*}(\tilde{\pi}).$$

*Με τον παραπάνω αλγόριθμο πετυχαίνεται ο καθορισμός της άριστης απόφασης* $\alpha^*$ *καθώς και το λειτουργικό* gradient $\gamma_{\alpha^*}(\tilde{\pi})$ της Ηυ *για το* ***συγκεκριμένο δ.π*** $\tilde{\pi}$ .

## 2.3. Προσδιορισμός του συνόλου $\Gamma_H$ των λειτουργικών gradient vectors για την συνάρτηση Ηυ.

Αποδεικνύεται ότι για δοσμένα, $\theta \in \Theta, \alpha \in A$ η συνάρτηση $l(.,\theta,\alpha)$ επάγει μία διαμέριση από κυρτές περιοχές (πολύεδρα) του χώρου Π, έτσι ώστε η $l(\pi,\theta,\alpha)$ να έχει μια μοναδική τιμή σε κάθε μια από αυτές. Η απόδειξη στηρίζεται στο γεγονός ότι η συνάρτηση $T(.,\theta,\alpha)$ μετασχηματίζει ευθύγραμμα τμήματα σε ευθύγραμμα τμήματα (βλέπε και πρόταση 1.4.1).

Συμβολίζουμε με $\boldsymbol{S}_{\alpha,\theta}$ την εν λόγω διαμέριση και $S^{l}_{a,\theta}$, $l$=1,2,…,k τις περιοχές που την αποτελούν:

$$\boldsymbol{S}_{\alpha,\theta} = \{ S^{1}_{a,\theta}, S^{2}_{a,\theta}, ..., S^{k}_{a,\theta} \}.$$

Ενδεχομένως κάποιες από τις παραπάνω περιοχές είναι κενές. Έτσι λοιπόν

$$\pi \in S^{l}_{a,\theta} \Rightarrow l(\pi,\theta,\alpha) = l.$$

Επομένως

$$S^{l}_{a,\theta} := \{\pi \in \Pi : T(\pi,\theta,\alpha).\gamma^{l} \geq T(\pi,\theta,\alpha).\gamma^{j},\ 1 \leq j \leq k\} =$$

$$= \{\pi \in \Pi : \pi.\, P^{\alpha}.R^{\alpha}_{\theta}.\gamma^{l} \geq \pi.\, P^{\alpha}.R^{\alpha}_{\theta}.\gamma^{j},\ 1 \leq j \leq k\} \qquad \textbf{2.3.1}$$

Για δοσμένο $\alpha \in A$, σχηματίζονται οι διαμερίσεις $\boldsymbol{S}_{\alpha,1}, \boldsymbol{S}_{\alpha,2}, \ldots, \boldsymbol{S}_{\alpha,M}$ του χώρου Π,(όπου M=$|\Theta|$).Μπορούμε να σχηματίσουμε μια διαμέριση γινόμενο από αυτές τις διαμερίσεις $\boldsymbol{S}_{\alpha,\theta}$ δηλαδή:

$$\boldsymbol{S}_{\alpha} = \otimes_{\theta \in \Theta} \boldsymbol{S}_{\alpha,\theta} = \boldsymbol{S}_{\alpha,1} \otimes \boldsymbol{S}_{\alpha,2} \otimes \ldots \otimes \boldsymbol{S}_{\alpha,M}$$

που αποτελείται από τομές της μορφής

$$\boldsymbol{S}^{s}_{\alpha} = \bigcap_{\theta=1}^{M} S^{l_\theta}_{\alpha,\theta} \qquad s=(l_1, l_2, ..., l_M) \quad \text{όπου} \quad 1 \leq l_\theta \leq k, \theta = 1,2,3,...M. \qquad \textbf{2.3.2}$$

Οι τομές αυτές αποτελούν κυρτές περιοχές (κυρτά πολύεδρα) του χώρου Π.

Σημειώνουμε επίσης ότι για όλα τα $\pi \Xi \text{int}\, S_a^s$ αντιστοιχεί σταθερό gradient $\gamma_\alpha(\pi)$,

$$\gamma_\alpha(\pi)=q^\alpha+\beta.P^\alpha.\sum_\theta R_\theta^\alpha\ \gamma^{l(\pi,\theta,\alpha)} = q^\alpha+\beta.P^\alpha.\sum_\theta R_\theta^\alpha\ \gamma^{l_\theta},$$

που το συμβολίζουμε με $\xi_a^s$.

Δηλαδή $$\gamma_\alpha(\pi)=\xi_a^s = q^\alpha+\beta.P^\alpha.\sum_\theta R_\theta^\alpha\ \gamma^{l_\theta}, \qquad \textbf{2.3.3}$$

όπου $$s=(l_1,l_2,\ldots,l_M).$$

Συμπεραίνουμε ότι για κάθε $\pi \Xi \text{int}\, (S_a^s)$, έχουμε:

$$H_\alpha \upsilon(\pi)=\pi.\xi_a^s$$

Πράγματι, $$H_\alpha \upsilon(\pi)=\pi.q^a+\beta.\sum_\Theta\{\theta/\pi,\alpha\}.\upsilon(\mathrm{T}(\pi,\theta,\alpha))$$

$$=\pi.(q^\alpha+\beta.P^\alpha.\sum_\theta R_\theta^\alpha\,.\,\gamma^{l(\pi,\alpha,\theta)})$$

$$=\pi.\gamma_\alpha(\pi)=\pi.\xi_a^s. \qquad \textbf{2.3.4}$$

Εστω τώρα $$G=\bigcup_a G^\alpha,$$

Όπου $G^\alpha=\{\zeta_a^S := q^\alpha+\beta.P^\alpha.\sum_\theta R_\theta^\alpha.\gamma^{j_\theta} : s=(j_1,j_2,\ldots,j_M), 1\le j_\theta\le k, 1\le\theta\le M\}, \alpha\in A$.

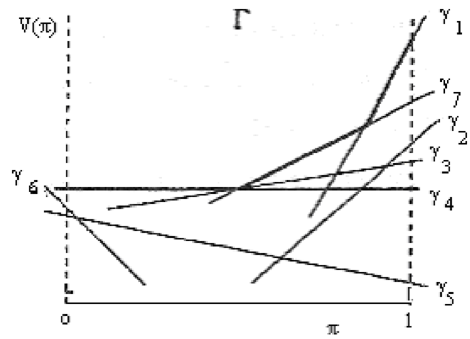
*Κάθε «gradient» του συνόλου $G^\alpha$, λέμε ότι έχει φέρουσα απόφαση την α.*

Σημειώνουμε ότι G είναι το σύνολο των ΄΄εν δυνάμει΄΄*«gradient-vectors» για την συνάρτηση* H υ

Τότε $$H\upsilon(\pi)=\max\{\pi\gamma:\gamma\in G\}, \pi\in\Pi. \qquad \textbf{2.3.5}$$

Οι πληθάριθμοι των συνόλων $G^\alpha$, $\alpha\in A$ και G είναι αντίστοιχα

$$|G^\alpha|=|\Gamma|^{|\Theta|}=k^M,\ \alpha\in A \text{ και } |G|=|A|.k^M.$$




**Σχήμα 2.1:Ένα παράδειγμα μιας P.W.L.C συνάρτησης τιμών με άχρηστα (useless-vectors) τα $\gamma_2,\gamma_3,\gamma_5,\gamma_6$ στην ελάχιστη αντιπροσώπευση της.**

Το παραπάνω σύνολο G ενδέχεται να είναι ένα πολυπληθές σύνολο. Αντικειμενικός μας σκοπός,είναι να οικοδομήσουμε ένα σύνολο $\Gamma_H$ ελάχιστο στην αντιπροσώπευση της H υ, δηλαδή ένα σύνολο:

$$\Gamma_H=\{\gamma: \gamma\in G \text{ και } R(\gamma,G)\neq \varnothing \} \qquad \textbf{2.3.6}$$

Τότε βέβαια: $[H\, \upsilon](\pi)=\max\{\pi\,\gamma: \gamma\in G\}$

$$=\max\{\pi\,\gamma: \gamma\in \Gamma_H\}. \qquad \textbf{2.3.7}$$

Με άλλα λόγια, ένα εν δυνάμει «gradient» είναι λειτουργικό και τοποθετείται στο σύνολο $\Gamma_H$, όταν η συσχετιζόμενη περιοχή $R(\gamma,G)\neq \varnothing$.

Η μέθοδος του γραμμικού προγραμματισμού μπορεί να χρησιμοποιηθεί για την περικοπή των μη λειτουργικών *«gradient-vectors» του συνόλου G με σκοπό να επιτύχουμε την ελάχιστη αντιπροσώπευση* **$\Gamma_H$** για την Ηυ.

Έστω $\gamma' \in G$ .Θεωρούμε το γραμμικό πρόγραμμα (γ.π)

$$z^* = \max z$$

υπό τους περιορισμούς

$$z \le \pi.(\gamma' - \gamma) \quad \forall \gamma \in G$$

$$\sum_{i=1}^{N} \pi_i = 1 \ , \ \ \pi_i \ge 0, \ \ 1 \le i \le N \, .$$

Το παραπάνω γ.π έχει N+1μεταβλητές (τις $z,\pi_1,\pi_2,\ldots,\pi_N$) και $|G|+N+1$ περιορισμούς.Το *gradient-vector* $\gamma'$ είναι λειτουργικό ($\gamma' \in \Gamma_H$) αν και μόνον αν $z^* = 0$.

Επομένως επιλύοντας το γ.π. τοποθετούμε το *$\gamma'$ στο σύνολο* $\Gamma_H$ αν $z^* = 0$.

Επαναλαμβάνουμε την ίδια διαδικασία για όλα τα *gradient-vectors του συνόλου G, πράγμα που συνεπάγεται την επίλυση* $|G|$ το πλήθος γ.π. για τον καθορισμό του συνόλου $\Gamma_H$. Επισημαίνουμε την υπολογιστική δυσκολία που οφείλεται στο μεγάλο μέγεθος του συνόλου G ($|G| = |A|.k^M$).

## 2.4. Προσδιορισμός της υποστηρίζουσας περιοχής R($\gamma^*$,$\Gamma_H$)

Εφαρμόζοντας τον αλγόριθμο του ενός βήματος της ενότητας 2.2 μπορούμε να βρούμε ένα συγκεκριμένο gradient ,$\gamma^*$,που είναι λειτουργικό στο $\tilde{\pi}$ και την αντίστοιχη άριστη απόφαση $\alpha^*$ , ( $\gamma^* = \gamma_{a^*}(\tilde{\pi})$)

Στην τρέχουσα ενότητα έχουμε σαν αντικείμενο να βρούμε την περιοχή $R(\gamma^*, \Gamma_H)$ ,όπου το συγκεκριμένο gradient ,$\gamma^*$,είναι λειτουργικό. Φανταζόμαστε προς τούτο ότι μετακινούμαστε σε ένα δ.π , π, λίγο πιο μακριά του $\tilde{\pi}$ ,και υπολογίζουμε με βάση τον αλγόριθμο του ενός βήματος το λειτουργικό gradient γ για το π. Εξακολουθούμε να μετακινούμαστε μέχρι να πετύχουμε $\gamma \neq \gamma^*$.

Σύμφωνα με τον Sondik [119] υποδεικνύονται δύο δυνατότητες να επιτευχθεί το παραπάνω $\gamma \neq \gamma^*$ μεταβαίνοντας από το $\tilde{\pi}$ στο π.

**i)** Να αλλάξει ο δείκτης $l(\tilde{\pi},\theta,\alpha^*)$ σε $l(\pi,\theta,\alpha^*) \neq l(\tilde{\pi},\theta,\alpha^*)$ για κάποιο $\theta \in \Theta$

ή

**ii)** Να αλλάξει η βέλτιστη απόφαση α* που αντιστοιχεί στο $\tilde{\pi}$.

Για δοσμένο θ, η περιοχή του χώρου Π για την οποία ισχύει $l(\tilde{\pi},\theta,\alpha^*) = l(\pi,\theta,\alpha^*)$ δίνεται από τη σχέση : $\pi.P^{\alpha^*}.R_\theta^{\alpha^*}.(\gamma^{l(\tilde{\pi},\theta,\alpha^\bullet)} - \gamma^j) \geq 0$ με $1 \leq j \leq k$.

Επαναλαμβάνοντας την ίδια διαδικασία για όλα τα θ=1,2,…,M η συνθήκη $l(\tilde{\pi},\theta,\alpha^*) = l(\pi,\theta,\alpha^*)$ , $1 \leq \theta \leq M$ ισχύει αν και μόνο αν :

$$\pi.P^{\alpha^*}.R_\theta^{\alpha^*}.(\gamma^{l(\tilde{\pi},\theta,\alpha^\bullet)} - \gamma^j) \geq 0 \quad \text{με } 1 \leq j \leq k, 1 \leq \theta \leq M \qquad \textbf{2.4.1}$$

Οι γραμμικές ανισότητες (2.4.1) ορίζουν την περιοχή που η δυνατότητα (i) αποκλείεται. Σύμφωνα πάντα με τον Sondik [119], η απόφαση $\alpha^*$ είναι βέλτιστη στα δ.π της συσχετιζόμενης περιοχής του $\gamma^*, R(\gamma^*, \Gamma_H)$ αν

$$\pi.(\gamma^*-\gamma(\alpha)) \geq 0 \qquad \forall\, \alpha \in A, \qquad \textbf{2.4.2}$$

όπου :

$$\gamma(\alpha) = q^\alpha + \beta.P^\alpha \sum_\theta R_\theta^\alpha \ \gamma^{l(\tilde{\pi},\theta,\alpha)}.$$

Σύμφωνα με τον παραπάνω συμβολισμό έχουμε ότι:

$$\gamma^* = \gamma(\alpha^*) = q^{\alpha^\bullet} + \beta.P^{\alpha^\bullet}.\sum_{\theta\in\theta} R_\theta^{\alpha^*}\ .\ \gamma^{l(\tilde{\pi},\theta,\alpha^\bullet)}.$$

*Ο Sondik [119] είχε θεωρήσει μόνο τις ανισότητες (2.4.1) και (2.4.2) για να προσδιορίσει την περιοχή $R(\gamma^*,\Gamma_H)$,το οποίο δεν είναι ορθό, όπως επισημάνθηκε από τον Lovejoy [77],διότι ενδέχεται σε κάποιο π,καθώς απομακρυνόμαστε από το $\tilde{\pi}$ να ικανοποιείται η ανίσωση:*

*$\pi.(\gamma^*-\gamma(\alpha)) \geq 0 \ \ \forall\, \alpha \in A$ και να έχουμε ωστόσο $l(\pi,\theta,\alpha) \neq l(\tilde{\pi},\theta,\alpha)$, για κάποιο $\theta \in \Theta$,*

*οπότε για το «gradient» $\gamma'(\alpha) = q^\alpha + \beta.P^\alpha \sum_\theta R_\theta^\alpha \ \gamma^{l(\pi,\theta,\alpha)}$ μπορεί να ισχύει η ανισότητα:*

$$\pi.\gamma'(\alpha) > \pi.\gamma^* > \pi.\gamma(\alpha) \quad \text{για κάποιο } \alpha \in A.$$

Ο *Lovejoy* [77],τεκμηρίωσε το παραπάνω σφάλμα με αριθμητικό παράδειγμα. Επομένως οι σχέσεις (2.4.1) και (2.4.2) δεν επαρκούν για τον προσδιορισμό της συσχετιζόμενης περιοχής $R(\gamma^*,\Gamma_H)$ *του $\gamma^*$. Σύμφωνα με τον Lovejoy,σε περίπτωση που δεν γνωρίζουμε το σύνολο $\Gamma_H$ των λειτουργικών «gradients» για την Ηυ, η περιοχή $R(\gamma^*,\Gamma_H)$ καθορίζεται από τα $\pi \in \Pi$ για τα οποία:*

$$\pi.(\gamma^*-\gamma) \geq 0 \quad \forall\, \gamma \in G. \qquad \textbf{2.4.3}$$

όπου G είναι το σύνολο των εν δυνάμει *«gradient vectors» για την Ηυ* (βλέπε ενότητα 2.3).Σημειώνουμε ότι οι ανισώσεις (2.4.1) και (2.4.2) υπάγονται στην (2.4.3).Γενικά μόνο ένα υποσύνολο των περιορισμών (2.4.3) είναι απαραίτητο για τον καθορισμό της *περιοχήs $R(\gamma^*,\Gamma_H)$.Για την εύρεση αυτού του* υποσυνόλου εφαρμόζουμε τη μέθοδο του γραμμικού προγραμματισμού.

Έστω $\gamma' \in G$ .Θεωρούμε το ακόλουθο γραμμικό πρόγραμμα (γ.π)

$$z = \min_{\pi} \pi.(\gamma^* - \gamma')$$

υπό τους περιορισμούς

$$\pi.(\gamma^* - \gamma) \geq 0 \ , \forall\, \gamma \in G.$$

$$\sum_{i=1}^{N} \pi_i = 1 \ ,$$

$$\pi_i \geq 0, \quad 1 \leq i \leq N \, .$$

Αυτό το γ.π έχει N μεταβλητές (τις $\pi_1, \pi_2, \ldots, \pi_N$) και $|G| + N + 1$ περιορισμούς. Αν z=0 τότε ο περιορισμός

$$\pi.(\gamma^* - \gamma') \geq 0 \qquad \textbf{2.4.4}$$

είναι απαραίτητος για τον καθορισμό τής *περιοχής $R(\gamma^*, \Gamma_H)$*. Αντιθέτως αν $z \succ 0$, ο περιορισμός (2.4.4) μπορεί να παραλειφθεί ως πλεονάζων.

Επαναλαμβάνουμε την ίδια διαδικασία για όλα τα *gradient-vectors του συνόλου G, πράγμα που συνεπάγεται την επίλυση* $|G|$ το πλήθος γ.π. προκειμένου να βρούμε το υποσύνολο των περιορισμών στην (2.4.3) που είναι απαραίτητοι για τον καθορισμό του συνόρου του κυρτού πολυέδρου *$R(\gamma^*, \Gamma_H)$.Στο σημείο* αυτό επισημαίνουμε την υπολογιστική δυσκολία που συνδέεται με το μέγεθος του συνόλου *G*, ($|G| = |A|.k^M$ ).

Αν το σύνολο $\Gamma_H$ των λειτουργικών gradients για την Ηυ είναι γνωστό (π.χ. με την μέθοδο που περιγράφηκε στην ενότητα 2.3), τότε από τον ορισμό 2.1.1 η περιοχή *$R(\gamma^*, \Gamma_H)$* καθορίζεται από τα $\pi \in \Pi$ για τα οποία

$$\pi.(\gamma^* - \gamma) \geq 0 \ , \forall\, \gamma \in \Gamma_H. \qquad \textbf{2.4.5}$$

Ομοίως στην περίπτωση αυτή εν γένει μόνο ένα υποσύνολο των περιορισμών (2.4.5) είναι απαραίτητο για τον καθορισμό του συνόρου της *περιοχής $R(\gamma^*, \Gamma_H)$.Τ*ο υποσύνολο αυτό προσδιορίζεται με παρόμοια διαδικασία επίλυσης $|\Gamma_H|$ το πλήθος γ.π. Τέλος επισημαίνουμε ότι η άριστη απόφαση για την περιοχή *$R(\gamma^*, \Gamma_H)$ είναι η φέρουσα απόφαση $\alpha^*$ του gradient-vector* $\gamma^*$. Πράγματι, για κάθε $\pi \in R(\gamma^*, \Gamma_H)$ έχουμε

$$H\upsilon(\pi) = \max\{\pi\, \gamma : \gamma \in \Gamma_H\} = \pi.\gamma^*.$$

# ΣΥΜΠΕΡΑΣΜΑΤΑ

Στο κεφάλαιο αυτό περιγράφουμε τη μέθοδο των Smallwood –Sondik- Lovejoy για τον υπολογισμό της συνάρτησης Hυ αν υ είναι γνωστή κατά τμήματα γραμμική και κυρτή συνάρτηση. Ο υπολογισμός της Hυ ανάγεται στον καθορισμό του συνόλου $\Gamma_H$ των λειτουργικών gradients της Hυ. Ειδικότερα:

- Παρουσιάζουμε τον αλγόριθμο του ενός βήματος των Smallwood –Sondik ,με τον οποίο υπολογίζεται η τιμή Hυ(π) για δοσμένο δ.π. π.
- Προσδιορίζουμε το σύνολο $\Gamma_H$ περικόπτοντας μη λειτουργικά gradients από το σύνολο G των εν δυνάμει gradient-vectors της Hυ επιλύοντας $|G|$ το πλήθος γραμμικά προγράμματα.
- Για κάθε λειτουργικό gradient-vector της Hυ προσδιορίζουμε το σύνορο της συσχετιζόμενης περιοχής με την επίλυση γραμμικών προγραμμάτων.
- Επισημαίνουμε τις υπολογιστικές δυσκολίες, που οφείλονται στο μεγάλο μέγεθος του συνόλου G.

Στην κατεύθυνση λοιπόν αναζήτησης μιας εναλλακτικής μεθόδου υπολογισμού της Hυ, παρακάμπτοντας το σύνολο G,προτείνεται ένας νέος αλγόριθμος στο επόμενο κεφάλαιο.