EE728
Προχωρημένα Θέματα Θεωρίας Πληροφορίας
6η διάλεξη
(2η έκδοση, 2/4/1011)

Δημήτρης-Αλέξανδρος Τουμπακάρης

Τμήμα ΗΜ&ΤΥ, Πανεπιστήμιο Πατρών

29 Μαρτίου 2011

Περιεχόμενα 6ης διάλεξης

Μεγιστοποίηση κοίλης συνάρτησης κατανομής πιθανότητας

- Θεωρούμε συνάρτηση $f(\mathbf{p}) : \mathbf{R}^n \to \mathbf{R}$ η οποία είναι κοίλη $\cap$ ως προς $\mathbf{p}$.
- Έστω, επίσης, ότι το $\mathbf{p}$ είναι κατανομή (διάνυσμα πιθανότητας), δηλαδή $p_i \geq 0,\ i = 1, \ldots, n$ και $\sum_{i=1}^{n} p_i = \mathbf{1}^T \mathbf{p} = 1$.
- Τέλος, θεωρούμε ότι οι μερικές παράγωγοι $\partial f(\mathbf{p})/\partial p_i$ ορίζονται και ότι είναι συνεχείς με μοναδική εξαίρεση το $\lim_{p_i \to 0} \partial f(\mathbf{p})/\partial p_i$ που μπορεί να είναι και $+\infty$.

Μεγιστοποίηση κοίλης συνάρτησης κατανομής πιθανότητας (συνέχεια)

- Αποδεικνύεται ότι οι παρακάτω συνθήκες είναι ικανές και αναγκαίες για να μεγιστοποιείται η $f()$ στο σημείο (κατανομή) $\mathbf{p}$.

$$\frac{\partial f(\mathbf{p})}{\partial p_i} = \lambda, \text{ για όλα τα } i \text{ για τα οποία } p_i > 0$$
$$\frac{\partial f(\mathbf{p})}{\partial p_i} \leq \lambda, \text{ για όλα τα } i \text{ για τα οποία } p_i = 0$$

  για κάποια τιμή της παραμέτρου λ.

- Για την απόδειξη δείτε π.χ. Gallager Theorem 4.4.1.

Μεγιστοποίηση αμοιβαίας πληροφορίας

- Με χρήση του προηγούμενου θεωρήματος και του ότι η $I(X;Y)$ είναι κοίλη $\cap$ συνάρτηση της κατανομής εισόδου $p(x)$ για δεδομένο κανάλι $p(y|x)$, αποδεικνύεται ότι οι παρακάτω δύο συνθήκες αποτελούν ικανή και αναγκαία συνθήκη για να επιτυγχάνει μια κατανομή $\mathbf{p}^*$ τη χωρητικότητα.

$$I(X = x_i; Y) = C, \text{ για όλα τα } x_i \text{ για τα οποία } p^*(x_i) > 0$$
$$I(X = x_i; Y) \leq C, \text{ για όλα τα } x_i \text{ για τα οποία } p^*(x_i) = 0$$

όπου $I(X = x_i; Y) = \sum_{y \in |\mathcal{Y}|} p(y|x_i) \log \frac{p(y|x_i)}{p(y)}$ η αμοιβαία πληροφορία μεταξύ $X = x_i$ και Y.

Μεγιστοποίηση αμοιβαίας πληροφορίας (συνέχεια)

- Το αποτέλεσμα αυτό έχει μια διαισθητική επεξήγηση: Εάν για $x_i \neq x_j$ $I(X = x_i; Y) > I(X = x_j; Y)$, μπορούμε να αυξήσουμε την $I(X; Y) = \sum_{x_k} p(x_k) I(X = x_k; Y)$ χρησιμοποιώντας τη x_i πιο συχνά και τη x_j λιγότερο συχνά (αλλάζοντας τις $p(x_i)$ και $p(x_j)$).
- Αυτό έχει ως αποτέλεσμα να αλλάξει η $p(y) = \sum_{x_k} p(x_k) p(y|x_k)$.
- Τελικά, η διαδικασία αυτή θα ισορροπήσει σε σημείο όπου όλες οι $I(X = x_i; Y)$ εκτός, ίσως, από κάποιες που αντιστοιχούν σε κακές εισόδους, θα ισούνται μεταξύ τους (και, επομένως, και με τη χωρητικότητα, C).

Άλλες ενδιαφέρουσες ιδιότητες και αποτελέσματα

- Αναφέρουμε, τέλος, 3 ενδιαφέροντα πορίσματα. Για αποδείξεις δείτε π.χ. Gallager Κεφ. 4.5.

- **Πόρισμα 6.1.** Για οποιαδήποτε κατανομή εισόδου, $p^*(x)$, που επιτυγχάνει τη χωρητικότητα σε διακριτό κανάλι χωρίς μνήμη, όλες οι πιθανότητες συμβόλων εξόδου, $p(y)$, είναι αυστηρώς θετικές (αρκεί για κάθε έξοδο να υπάρχει τουλάχιστον μία είσοδος που οδηγεί σε αυτήν).

- **Πόρισμα 6.2.** Η κατανομή εξόδου, $p^*(y)$, για την οποία $I(X;Y) = C$ είναι μοναδική. Όλες οι κατανομές εισόδου, $p(x)$, για τις οποίες $\sum_{x\in|\mathcal{X}|} p(x)p(y|x) = p^*(y)$ επιτυγχάνουν τη χωρητικότητα.

- **Πόρισμα 6.3.** Έστω m ο *ελάχιστος* αριθμός συμβόλων εισόδου που μπορούν να χρησιμοποιηθούν (με μη μηδενική πιθανότητα) για να επιτευχθεί μετάδοση με τη χωρητικότητα. Έστω $\mathcal{A}$ ένα τέτοιο σύνολο m συμβόλων εισόδου. Ισχύει $m \leq |\mathcal{Y}|$. Επίσης, η κατανομή $p(x)$ στα στοιχεία του $\mathcal{A}$ που επιτυγχάνει τη χωρητικότητα είναι μοναδική.

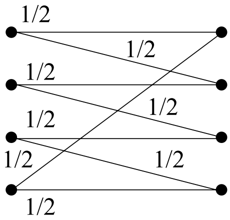
Πώς υπολογίζουμε τη χωρητικότητα;

- Γενικά, ο υπολογισμός της χωρητικότητας δεν είναι εύκολη υπόθεση.
- Σε μερικές, ειδικές, περιπτώσεις μπορούμε να χρησιμοποιήσουμε ιδιότητες όπως, π.χ. στην περίπτωση συμμετρικών καναλιών.
- Άλλες φορές μπορούμε να "μαντέψουμε" την κατανομή εισόδου και να δείξουμε ότι επιτυγχάνει ένα άνω φράγμα για τη χωρητικότητα (όπως κάναμε για το συμμετρικό κανάλι).
- Στη γενική περίπτωση καταφεύγουμε σε αριθμητικές μεθόδους με χρήση υπολογιστή. Μια ευρέως χρησιμοποιούμενη μέθοδος είναι των Blahut & Arimoto. Τα τελευταία χρόνια έχουν προταθεί βελτιώσεις που συγκλίνουν πολύ πιο γρήγορα σε σχέση με τον αρχικό αλγόριθμο.

Χωρητικότητα καναλιών με ανάδραση

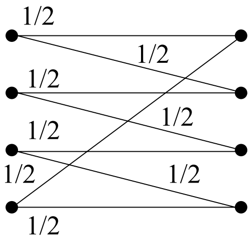
1. Παρατηρήσεις και θεωρήματα σχετικά με τη χωρητικότητα
2. Χωρητικότητα καναλιών με ανάδραση (feedback)
   - Εισαγωγή, Ορισμοί και Μοντέλο
   - $C_{FB} = C$
3. Αποκωδικοποίηση Μέγιστης Πιθανοφάνειας και Εκθέτης Σφάλματος
   - Αποκωδικοποίηση Μέγιστης Πιθανοφάνειας
   - Εκθέτης Σφάλματος

Παράδειγμα 6.1



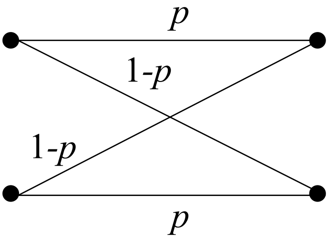
- Θεωρούμε το διακριτό κανάλι χωρίς μνήμη του σχήματος ("ενθόρυβη γραφομηχανή").
- Η χωρητικότητα του καναλιού ισούται με $C = \max I(X;Y) = \max\{H(Y) - H(Y|X)\} = 2 - 1 = 1$ bit.
- Μπορούμε να επιτύχουμε μετάδοση με ρυθμό ίσο με τη χωρητικότητα και με μηδενική πιθανότητα σφάλματος χρησιμοποιώντας π.χ. τις εισόδους 0 και 2. Προφανώς, $R = 1$ bit $= C$.

Παράδειγμα 6.1 (συνέχεια)



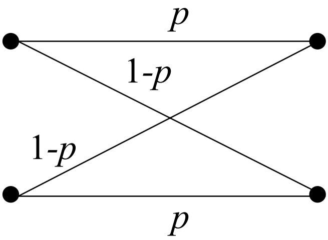
- Ό,τι και να συμβεί στο κανάλι είμαστε βέβαιοι ότι δε θα εμφανιστεί σφάλμα αποκωδικοποίησης.
- Εάν μπορούσαμε να χρησιμοποιήσουμε ανάδραση (feedback), η χωρητικότητα θα παρέμενε η ίδια;

Παράδειγμα 6.2



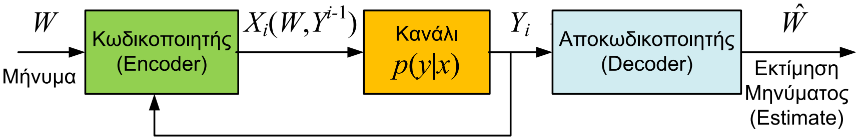
- Ας θεωρήσουμε, τώρα, το δυαδικό συμμετρικό κανάλι.
- Γνωρίζουμε ότι $C = 1 - H(p)$ και ότι η χωρητικότητα επιτυγχάνεται χρησιμοποιώντας και τα δύο μηνύματα με ίση πιθανότητα. Επομένως, κάθε φορά που στέλνουμε ένα από τα δύο μηνύματα στο κανάλι δε γνωρίζουμε εάν το μήνυμα μεταδόθηκε επιτυχώς. Η πιθανότητα σφάλματος <u>ανά μετάδοση</u> είναι μη μηδενική.

Παράδειγμα 6.2 (συνέχεια)



- Τι συμβαίνει εάν χρησιμοποιήσουμε ανάδραση; (όπου γνωρίζουμε εάν έχει εμφανιστεί σφάλμα στο δέκτη;)
- Σημείωση: Όταν χρησιμοποιούμε ανάδραση στο BSC, ο πομπός γνωρίζει ότι συνέβη σφάλμα, όχι, όμως, ο δέκτης!
- Παρόλο που κανείς θα περίμενε, ίσως, το αντίθετο, θα αποδείξουμε ότι, σε διακριτά κανάλια χωρίς μνήμη, η χρήση ανάδρασης δεν αυξάνει τη χωρητικότητα!

Χωρητικότητα καναλιού με ανάδραση – Μοντέλο



- Στο μοντέλο του σχήματος θεωρούμε ότι ο δέκτης στέλνει όλα τα ληφθέντα σύμβολα Y_i στον πομπό άμεσα και χωρίς σφάλματα. Ο πομπός χρησιμοποιεί την πληροφορία που λαμβάνει από το δέκτη προκειμένου να αποφασίσει πώς θα μεταδώσει.

Χωρητικότητα καναλιού με ανάδραση – Ορισμοί

- **Ορισμός 6.4.** Κώδικας ανάδρασης (feedback code) $(2^{nR}, n)$:
  - Ένας κωδικοποιητής που παράγει ακολουθία $x_i(W, Y^{i-1})$, όπου κάθε x_i είναι συνάρτηση του τρέχοντος μηνύματος W, καθώς και των σημάτων που ελήφθησαν στο δέκτη έως και τη χρονική στιγμή $i-1$: $Y_1, Y_2, \ldots, Y_{i-1}$ και
  - Ένας αποκωδικοποιητής $g : \mathcal{Y}^n \to \{1, 2, \ldots, 2^{nR}\}$.
- Θεωρούμε ότι τα μηνύματα W είναι ομοιόμορφα κατανεμημένα. Επομένως, $P_e^{(n)} = \Pr\{g(Y^n) \neq W\}$, όπου $X^n = X^n(W)$.
- **Ορισμός 6.5.** Η (λειτουργική) χωρητικότητα με ανάδραση (feedback capacity), C_{FB}, του διακριτού καναλιού χωρίς μνήμη ισούται με το μέγιστο ρυθμό που είναι εφικτός με χρήση κωδίκων ανάδρασης.

Χωρητικότητα καναλιού με ανάδραση

- Θεώρημα 6.6. (Cover 7.12.1): $C_{FB} = C = \max_{p(x)} I(X;Y)$.
- Απόδειξη Είναι, κατ' αρχάς, προφανές ότι $C_{FB} \geq C$ (ευθύ), δεδομένου ότι το κανάλι χωρίς ανάδραση μπορεί να θεωρηθεί ως ειδική περίπτωση του καναλιού με ανάδραση.
- Θα αποδείξουμε ότι $C \geq C_{FB}$ και, επομένως, $C = C_{FB}$.
- Θα χρησιμοποιήσουμε και πάλι την ανισότητα Fano, όπως και στο αντίστροφο του Θεωρήματος Κωδικοποίησης Καναλιού. Ωστόσο, στην απόδειξη πρέπει να ληφθεί υπόψη ότι στο κανάλι με ανάδραση <u>δεν ισχύει</u> η σχέση $I(X^n;Y^n) \leq nC$.

Χωρητικότητα καναλιού με ανάδραση (2)

- Αρκεί να χρησιμοποιήσουμε την εναλλακτική απόδειξη του αντιστρόφου της προηγούμενης εβδομάδας.
- Η απόδειξη επαναλαμβάνεται αυτούσια για διευκόλυνση και για να τονιστεί ότι δεν επηρεάζεται από την ύπαρξη ή μη ανάδρασης.
- Επειδή $W \to Y^n \to \hat{W}$,

$$H(W|Y^n) \leq H(W|\hat{W}).$$

- Επίσης, από την ανισότητα Fano,

$$H(W|\hat{W}) \leq 1 + P_e^{(n)} nR, \text{ όπου } P_e^{(n)} = \Pr\{W \neq \hat{W}\}.$$

Χωρητικότητα καναλιού με ανάδραση (3)

- Υποθέτοντας, και πάλι, ότι η τ.μ. W ακολουθεί ομοιόμορφη κατανομή,

$$\begin{aligned} nR = H(W) &\overset{(a)}{=} I(W;Y^n) + H(W|Y^n) \\ &\overset{(b)}{\leq} I(W;Y^n) + 1 + P_e^{(n)} nR \end{aligned}$$

(a) Σχέση εντροπίας-αμοιβαίας πληροφορίας, (b) ανισότητα Fano.

Χωρητικότητα καναλιού με ανάδραση (4)

$$
\begin{aligned}
nR &\leq I(W;Y^n) + 1 + P_e^{(n)} nR \\
&\overset{(c)}{=} \sum_{i=1}^{n} I(W;Y_i|Y^{i-1}) + 1 + P_e^{(n)} nR \\
&\overset{(d)}{\leq} \sum_{i=1}^{n} I(W,Y^{i-1};Y_i) + 1 + P_e^{(n)} nR \\
&\overset{(e)}{\leq} \sum_{i=1}^{n} I(X_i,W,Y^{i-1};Y_i) + 1 + P_e^{(n)} nR
\end{aligned}
$$

(c) κανόνας αλυσίδας, (d)
$I(W,Y^{i-1};Y_i) = I(Y^{i-1};Y_i) + I(W;Y_i|Y^{i-1})$, (e) $X_i = f(W,Y^{i-1})$.

Χωρητικότητα καναλιού με ανάδραση (5)

$$
\begin{aligned}
nR &\leq \sum_{i=1}^{n} I(X_i, W, Y^{i-1}; Y_i) + 1 + P_e^{(n)} nR \\
&\stackrel{(f)}{=} \sum_{i=1}^{n} I(X_i; Y_i) + 1 + P_e^{(n)} nR \\
&\leq nC + 1 + P_e^{(n)} nR = n\left(C + \frac{1}{n} + P_e^{(n)}\right)
\end{aligned}
$$

(f) $X_i = f(W, Y^{i-1})$ και το κανάλι δεν έχει μνήμη, οπότε $(W, Y^{i-1}) \rightarrow X_i \rightarrow Y_i$.

Χωρητικότητα καναλιού με ανάδραση (6)

- Επομένως, $I(W;Y^n) \leq nC$, και

$$nR \leq 1 + P_e^{(n)} nR + I(W;Y^n) \leq P_e^{(n)} nR + 1 + nC.$$

- Διαιρώντας με n, και για $n \to \infty$,

$$R \leq C, \text{ και, επομένως, } C_{FB} \leq C.$$

- Παρόλο που η χρήση ανάδρασης σε διακριτά κανάλια χωρίς μνήμη δεν αυξάνει τη χωρητικότητα, ενδέχεται να διευκολύνει τη μετάδοση. Για παράδειγμα, στο κανάλι διαγραφής, η μετάδοση απλουστεύεται εάν γνωρίζουμε πότε το σήμα εισόδου διαγράφεται.
- Φυσικά, στην πράξη, μπορεί να μην υπάρχει αξιόπιστος δίαυλος ανάδρασης ή να έχει κόστος (π.χ. σε εύρος ζώνης ή καθυστέρηση).

Αποκωδικοποίηση Μέγιστης Πιθανοφάνειας και Εκθέτης Σφάλματος

1. Παρατηρήσεις και θεωρήματα σχετικά με τη χωρητικότητα

2. Χωρητικότητα καναλιών με ανάδραση (feedback)
   - Εισαγωγή, Ορισμοί και Μοντέλο
   - $C_{FB} = C$

3. Αποκωδικοποίηση Μέγιστης Πιθανοφάνειας και Εκθέτης Σφάλματος
   - Αποκωδικοποίηση Μέγιστης Πιθανοφάνειας
   - Εκθέτης Σφάλματος

Αποκωδικοποίηση Μέγιστης εκ των Υστέρων Πιθανότητας (Maximum A Posteriori Probability - MAP)

- Για την απόδειξη του Θεωρήματος Κωδικοποίησης Καναλιού υποθέσαμε ότι η αποκωδικοποίηση βασίζεται στην Ιδιότητα Από Κοινού Ασυμπτωτικής Ισοδιαμέρισης (Joint AEP).
- Δείξαμε ότι εάν η αποκωδικοποίηση βασίζεται στο Joint AEP μπορούμε να μεταδώσουμε με ρυθμούς αυθαίρετα κοντά στη χωρητικότητα με αυθαίρετα μικρή πιθανότητα σφάλματος.
- Αποδείξαμε ότι δεν μπορούμε να υπερβούμε τη χωρητικότητα. Επομένως, η αποκωδικοποίηση με χρήση από κοινού τυπικών ακολουθιών είναι *ασυμπτωτικά βέλτιστη*.

Αποκωδικοποίηση Μέγιστης εκ των Υστέρων Πιθανότητας (2)

- Εάν το κριτήριο είναι να ελαχιστοποιηθεί η πιθανότητα σφάλματος στο δέκτη, πρέπει να χρησιμοποιηθεί αποκωδικοποίηση Μέγιστης εκ των Υστέρων Πιθανότητας (Maximum a Posteriori (MAP) probability decoding).
- Θεωρούμε την πιθανότητα $p(y^n|x^n(w))$ να ληφθεί η ακολουθία y^n στο δέκτη δεδομένου ότι εστάλη ακολουθία $x^n(w)$ η οποία αντιστοιχεί στο μήνυμα w (η κωδική λέξη του μηνύματος w).
- Από το Θεώρημα Ολικής πιθανότητας, $\Pr\{\hat{w} = w\}(= 1 - \Pr\{\hat{w} \neq w\}) = \sum_{y^n=1}^{|\mathcal{Y}|^n} p(y^n) \Pr\{\hat{w} = w|y^n\} = \sum_{y^n=1}^{|\mathcal{Y}|^n} p(y^n) \Pr\{\hat{x}^n = x^n|y^n\}$.
- Επειδή $p(y^n) \geq 0$, για να μεγιστοποιήσουμε την $\Pr\{\hat{w} = w\}$ αρκεί να μεγιστοποιήσουμε κάθε όρο $\Pr\{\hat{x}^n = x^n|y^n\}$.

Αποκωδικοποίηση Μέγιστης εκ των Υστέρων Πιθανότητας (3)

- Από τον κανόνα του Bayes,

$$p(w|y^n) = \frac{p(y^n|x^n(w))p(w)}{p(y^n)},$$

$$\text{όπου } p(y^n) = \sum_{w=1}^{|\mathcal{W}|} p(w)p(y^n|x^n(w)).$$

- Για να ελαχιστοποιηθεί η πιθανότητα σφάλματος, πρέπει να επιλεγεί το μήνυμα w το οποίο μεγιστοποιεί την εκ των υστέρων (a posteriori) πιθανότητα του w δεδομένης της ληφθείσας ακολουθίας y^n ($p(w|y^n)$).

Κανόνας αποκωδικοποίησης MAP

Κανόνας αποκωδικοποίησης MAP

$\hat{w} = g(y^n)$, τέτοιο ώστε

$$p(\hat{w}|y^n) \geq p(w'|y^n), \text{ για όλα τα } w' \neq \hat{w}, \ \hat{w}, w' \in \mathcal{W}$$

Εναλλακτική έκφραση

$$\hat{w} = g(y^n) = \arg\max_{w'} p(w'|y^n)$$

Αποκωδικοποίηση Μέγιστης εκ των Υστέρων Πιθανότητας (4)

- Με χρήση του κανόνα του Bayes,

$$p(w|y^n) \geq p(w'|y^n) \Rightarrow$$
$$\frac{p(y^n|x^n(w))p(w)}{p(y^n)} \geq \frac{p(y^n|x^n(w'))p(w')}{p(y^n)}$$

- Επομένως, ο κανόνας MAP μπορεί να γραφτεί ως:

Κανόνας MAP

$$p(y^n|x^n(w))p(w) \geq p(y^n|x^n(w'))p(w')$$

- Για κανάλι χωρίς μνήμη,

Κανόνας MAP για κανάλι χωρίς μνήμη

$$p(y^n|x^n(w)) = \prod_{i=1}^{n} p(y_i|x_i(w)).$$

Αποκωδικοποίηση Μέγιστης Πιθανοφάνειας (Maximum Likelihood (ML) decoding)

- Με βάση τον κανόνα MAP επιλέγεται το μήνυμα που ικανοποιεί τη σχέση $p(y^n|x^n(w))p(w) \geq p(y^n|x^n(w'))p(w')$ για όλα τα $w' \neq w$.
- Εάν όλα τα μηνύματα εκπέμπονται με την ίδια πιθανότητα (ομοιόμορφα), ο αποκωδικοποιητής μπορεί να αποκωδικοποιήσει με βάση τη σχέση

$$p(y^n|x^n(w)) \geq p(y^n|x^n(w')) \text{ για όλα τα } w' \neq w.$$

- Η αποκωδικοποίηση με βάση την παραπάνω σχέση ονομάζεται μέγιστης πιθανοφάνειας (Maximum Likelihood - ML). Στη γενική περίπτωση (όπου τα μηνύματα δεν ακολουθούν ομοιόμορφη κατανομή) δε μεγιστοποιεί την πιθανότητα να έχει μεταδοθεί το μήνυμα w δεδομένης της ακολουθίας y^n.
- Ωστόσο, μεγιστοποιείται η πιθανότητα να έχει ληφθεί η y^n δεδομένου του w.

Γιατί ML και όχι MAP;

- Στη γενική περίπτωση (όπου η κατανομή των μηνυμάτων στην είσοδο του καναλιού δεν είναι ομοιόμορφη) η αποκωδικοποίηση ML δεν είναι βέλτιστη.
- Ωστόσο, στην πράξη, η αποκωδικοποίηση ML χρησιμοποιείται συχνότερα από την αποκωδικοποίηση MAP. Κάποιοι από τους λόγους είναι οι εξής:
  - Πολύ συχνά, τα μηνύματα που στέλνονται είναι ισοπίθανα (π.χ. όταν έχει γίνει καλή συμπίεση πριν από τη μετάδοση), οπότε η αποκωδικοποίηση ML είναι βέλτιστη.
  - Αποδεικνύεται (βλ. π.χ. Cioffi, http://www.stanford.edu/group/cioffi/book/chap1.pdf) ότι, εάν η κατανομή των μηνυμάτων $p(w)$ είναι άγνωστη, η αποκωδικοποίηση ML ελαχιστοποιεί την πιθανότητα σφάλματος για τη "χειρότερη" κατανομή εισόδου.
- Πολλές φορές η αποκωδικοποίηση ML είναι πολύπλοκη, οπότε χρησιμοποιούνται υποβέλτιστες μέθοδοι. Περισσότερα στα μαθήματα Ψηφιακών Επικοινωνιών.

Εκθέτης Σφάλματος

Εκθέτης Σφάλματος (Error Exponent) (εισαγωγή)

- Σύμφωνα με το Θεώρημα Κωδικοποίησης Καναλιού, είναι δυνατόν να μεταδώσουμε σε διακριτό κανάλι χωρίς μνήμη με αυθαίρετα μικρή πιθανότητα σφάλματος, αρκεί ο ρυθμός μετάδοσης να μην υπερβαίνει τη χωρητικότητα.
- Αντίστροφα, δεν υπάρχει κώδικας με αυθαίρετα μικρή πιθανότητα σφάλματος ο οποίος επιτυγχάνει μετάδοση με ρυθμό μεγαλύτερο από τη χωρητικότητα καναλιού.
- Αποδείξαμε το Θεώρημα Κωδικοποίησης Καναλιού όταν ο δέκτης αποκωδικοποιεί με βάση την Ιδιότητα Από Κοινού Ασυμπτωτικής Ισο-διαμέρισης. Το Θεώρημα αποδεικνύεται και για αποκωδικοποίηση μέγιστης πιθανοφάνειας (ML – βλ. π.χ. Gallager).

Εκθέτης Σφάλματος (Error Exponent) (2)

- Στην απόδειξη, για να επιτύχουμε αυθαίρετα μικρή πιθανότητα σφάλματος, αφήσαμε το n να τείνει στο άπειρο.
- Τι συμβαίνει όταν το n είναι πεπερασμένο; Πώς μεταβάλλεται η πιθανότητα σφάλματος ως συνάρτηση του n;
- Ένας τρόπος να ποσοτικοποιηθεί η εξάρτηση της μέσης πιθανότητας σφάλματος από το n είναι ο εκθέτης σφάλματος (error exponent) ο οποίος παρέχει ένα άνω φράγμα όταν χρησιμοποιειται αποκωδικοποίηση μέγιστης πιθανοφάνειας.

Εκθέτης Σφάλματος (Error Exponent) (3)

- Θεώρημα 6.7. (Gallager 5.6.2 & Corollary 1): Έστω διακριτό κανάλι χωρίς μνήμη με πίνακα μετάβασης $p(y_j|x_k), j = 1, \ldots, J$ και $k = 1, \ldots, K$. Για δεδομένο n και R θεωρούμε το σύνολο των κωδίκων $(2^{nR}, n)$ των οποίων τα σύμβολα επιλέγονται ανεξάρτητα με βάση κατανομή $p(x)$. Εάν ο δέκτης χρησιμοποιεί αποκωδικοποίηση μέγιστης πιθανοφάνειας, για τη μέση τιμή σφάλματος υπολογισμένη για όλους τους τυχαίους κώδικες οι οποίοι παράγονται με βάση κατανομή $p^*(x)$ και για όλα τα πιθανά μηνύματα, ισχύει

$$P_e^{(n)} \leq \exp\{-nE_r(R)\},$$

όπου $E_r(R)$ είναι ο εκθέτης τυχαίας κωδικοποίησης ή εκθέτης σφάλματος (random coding/error exponent)

$$E_r(R) = \max_{0\leq\rho\leq 1} \max_{p(x)} \{E_0(\rho, p(x)) - \rho R\},$$

$p^*(x)$ η κατανομή που επιτυγχάνει τον $E_r(R)$ και

$$E_0(\rho, p(x)) = -\log \sum_{j=1}^{J} \left[\sum_{k=1}^{K} p(x_k) p(y_j|x_k)^{1/(1+\rho)} \right]^{1+\rho}.$$

Εκθέτης Σφάλματος (Error Exponent) (4)

- Παρόλο που η έκφραση για τον εκθέτη σφάλματος είναι σχετικά πολύπλοκη, βασίζεται σε απλά βήματα (βλ. Gallager 5.6).
- Εάν μπορούμε να υπολογίσουμε τον $E_r(R)$ για δεδομένο διακριτό κανάλι χωρίς μνήμη, αποκτούμε ένα φράγμα για την πιθανότητα σφάλματος για δεδομένο ρυθμό μετάδοσης και δεδομένο μήκος κώδικα n: $P_e^{(n)} \leq \exp\{-nE_r(R)\}$.
- Αποδεικνύεται ότι, για $0 \leq R < C$, $E_r(R) > 0$ και, επομένως, με κατάλληλη κωδικοποίηση, η πιθανότητα σφάλματος μπορεί να κρατηθεί αυθαίρετα κοντά στο μηδέν με χρήση κωδίκων κατάλληλου μήκους n.
- Όπως και στην περίπτωση αποκωδικοποίησης με χρήση από κοινού τυπικότητας, το γεγονός ότι $P_e^{(n)} \leq \exp\{-nE_r(R)\}$ δε συνεπάγεται ότι η πιθανότητα σφάλματος $P_{e,w}^{(n)}$ που αντιστοιχεί στην κωδική λέξη $x^n(w)$ θα είναι $\leq \exp\{-nE_r(R)\}$. Ωστόσο, αποδεικνύεται (Gallager 5.6 Corollary 2) ότι υπάρχει κώδικας $(2^{nR}, n)$ τέτοιος ώστε $P_{e,w}^{(n)} \leq 4\exp\{-nE_r(R)\}$ για όλα τα $w = 1, 2, \ldots, 2^{nR}$.