2011

Equations for Random Codes in Communication Course

Lecturer: Haim Permuter *Scribe: Moti Kadosh, Ran Badanes*

Entropy	$H(X) \triangleq \mathbb{E}[-\log_2 P_X(X)] = -\sum_{x\in\mathcal{X}} P_X(x)\log_2 P_X(x)$
Differential Entropy	$h(X) \triangleq -\int f_X(x)\log_2(f_X(x))\,dx \triangleq \mathbb{E}[-\log_2 f_X]$
Gaussian distribution	$h(X) = (1/2)\log_2 2\pi e\sigma^2\ ,\quad X\sim N\left(0,\sigma^2\right)$
Gaussian vector	$h(X_i) = (n/2)\log_2 2\pi e\|K\|^{1/n}\ ,\quad X_i\quad i.i.d \sim N\left(0,\sigma^2\right)\ ,\ \|K\| \le \prod_{i=1}^n K_{ii}$
Joint Entropy	$H(X,Y) \triangleq \mathbb{E}[-\log P(X,Y)] = -\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} P(x,y)\log P(x,y)$
Conditional Entropy	$H(X\|Y) \triangleq \mathbb{E}[-\log P(X\|Y)] = -\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} P(x,y)\log P(x\|y)$
	$H(X\|Y=y) = H(X\|y) \triangleq -\sum_{x\in\mathcal{X}} P(x\|y)\log P(x\|y)$
	$h\left(X\|Y\right) \triangleq -\int f_{X,Y}\left(x,y\right)\log_2 f_{X\|Y}\left(x\|y\right)\ dx\,dy = \mathbb{E}\left[-\log_2 f_{X\|Y}\left(X\|Y\right)\right]$
Entropy chain rule	$H(X,Y) = H(X) + H(Y\|X)\ ,\quad H(X^n) = \sum_{i=1}^n H(X_i\|X^{i-1})$
Prob. chain rule	$P(x^n) = \prod_{i=1}^n P(x_i\|x^{i-1})$
Mutual information	$I(X;Y) \triangleq \sum_{x,y} P(x,y)\log\frac{P(x,y)}{P(x)P(y)} = E\left[\log\frac{P(X\|Y)}{P(X)}\right]$
	$I\left(X;Y\right) = \sup_{\mathcal{Q},\mathcal{P}} I\left([X]_{\mathcal{P}};[Y]_{\mathcal{Q}}\right)$
Information chain rule	$I(X^n;Y) = \sum_{i=1}^n I(X_i;Y\|X^{i-1})$
Identities	$I(X;Y) = H(X) - H(X\|Y) = H(Y) - H(Y\|X)$
	$I(X;Y) = H(Y) + H(X) - H(Y,X)\ ,\quad I(Y;X) = I(X;Y)$
Relative entropy	$D(P_X\|\|Q_X) \triangleq \sum_x P(x)\log\frac{P(x)}{Q(x)} = \mathbb{E}_P\left[\log\frac{P(X)}{Q(X)}\right]$
	$I(X;Y) = D(P_{X,Y}\|\|P_X P_Y)$
Divergence properties	$D\left(f_X\\|g_X\right) \ge 0,\quad D\left(f_X\\|g_X\right) = 0 \Leftrightarrow f_X(x) = g_X(x)$
Convex function	$f(\lambda x_1 + \bar{\lambda}x_2) \le \lambda f(x_1) + \bar{\lambda}f(x_2)\quad \forall\ x_1,\ x_2$ in its domain $\forall\lambda\in[0,1]$
	$\frac{d^2 f(x)}{dx^2} \ge 0 \Longleftrightarrow f(x)$ is convex.
Examples	$D(P\|\|Q)$ - convex, $H(X)$ - concave, $I(X,Y)$ - concave in P_X for a fixed $P_{Y\|X}$
Jensen's inequality	$\mathbb{E}[f(x)] \ge f(\mathbb{E}[x])$ for a convex function $f(x)$
Inequalities	$D(P\|\|Q) \ge 0\ ,\quad I(X;Y) \ge 0\ ,\quad H(X) \le \log\|\mathcal{X}\|\ ,\quad H(X\|Y) \le H(X)$
	$H(X) \ge H(g(X)),\ H(X\|Y) \le H(X\|g(Y))$
Log sum	$\sum_{i=1}^n a_i\log\left(\frac{a_i}{b_i}\right) \ge \left(\sum_{i=1}^n a_i\right)\log\left(\frac{\sum_{i=1}^n a_i}{\sum_{i=1}^n b_i}\right)$
Kraft Inequelity	$\sum_i 2^{-l_i} \le 1\ ,\quad l_i$ - codeword length
Shannon-Fano	$l_i = \left\lceil\log\frac{1}{p_i}\right\rceil\ ,\quad l_i$ - codeword length
Data Processing	$X - Y - Z \Rightarrow I(X;Y) \ge I(X;Z)$
L.L.N	$\frac{1}{n}\sum_{i=1}^n X_i \xrightarrow[n\to\infty]{P} E[X]\ ,\ \forall\epsilon>0\quad \lim_{n\to\infty} Pr\left(\left\|\frac{1}{n}\sum_{i=1}^n X_i - E[X]\right\| > \epsilon\right) = 0$

Markov Chain	If $X \to Y \to Z$, then $X \leftarrow Y \leftarrow Z$ $P(x,y,z) = P(x)P(y\|x)P(z\|y)$, $P(z\|x,y) = P(z\|y)$, $P(x\|y,z) = P(x\|y)$
Markov Inequality	$\Pr(X \geq a) \leq \frac{E[X]}{a}$ for $X \geq 0$ and $a > 0$
Fano inequality	$H(X\|\hat{X}) \leq 1 + \epsilon \log \|\mathcal{X}\|$, Where $\Pr(\hat{X} \neq X) = \epsilon$
Capacity	C=$\max_{P_X} I(X;Y)$
No feedback	$P(x_i\|x^{i-1}, y^{i-1}) = P(x_i\|x^{i-1})$
Memoryless channel	$P(y_i\|y^{i-1}, x^i) = P(y_i\|x_i)$ (generally) $P(y^n\|x^n) = \prod_{i=1}^n P(y_i\|x_i)$ (without feedback)
Channel Coding with Side Information	Only Decoder: $C = \max_{p(x)} I(X;Y\|S) = I(X;Y,S)$ Both Encoder and Decoder: $C = max_{p(x\|s)} I(X;Y\|S)$
Known Capacities	BSC: $C = 1 - H(\delta)$ Erasure Channel: $C = 1 - \delta$, δ - Prob. for error m parallel Channels: $C = \log_2(\sum_{i=1}^m 2^{c_i})$, one usage at a time
Source coding	$R_X \leq H(X) + 1$, R_X-Minimum rate to compress i.i.d source X losslessly
Channel seperation	$H(v) \leq R \leq C$, v - Input vector, c - Channel capacity
Gaussian Channel	$Y_i = X_i + Z_i$, $Z_i \sim \mathcal{N}(0, \sigma_z^2)$, $C = \max_{f(x):E(X^2) \leq P} I(X;Y)$ $I(X;Y) \leq \frac{1}{2}\log_2(\frac{\sigma_z^2 + P}{\sigma_z^2}) = \frac{1}{2}\log_2(1 + SNR)$, equality if $X \sim N(0,P)$
Volume of the set	$Vol\,(A^n) = \int_{X^n \in A^n} dx_1 dx_2 \ldots dx_n$,X continuous
Weak typicality	$A_\epsilon^{(n)}(x,y) = \{(x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n : \|-\frac{1}{n}\log p(x^n) - H(X)\| \leq \epsilon$ $\|-\frac{1}{n}\log p(y^n) - H(Y)\| \leq \epsilon$ $\|-\frac{1}{n}\log p(x^n, y^n) - H(X,Y)\| \leq \epsilon\}$
Typical set properties	$2^{-(nH(X,Y)+\epsilon)} \leq \Pr(x^n, y^n) \leq 2^{-(nH(X,Y)-\epsilon)}$ x^n, y^n i.i.d $\Pr((X^n, Y^n) \in A_\epsilon^n(x^n, y^n)) \xrightarrow[n \to \infty]{} 1$ x^n, y^n i.i.d $Vol\,(A_\epsilon^n(x^n, y^n)) \leq 2^{(nh(x,y)+\epsilon)}$ $Vol\,(A_\epsilon^n(x^n, y^n)) \geq (1-\epsilon)2^{(nh(x,y)-\epsilon)}$ $H(X) - \epsilon \leq \frac{1}{n}\log\|\mathcal{A}_\epsilon^{(n)}\| \leq H(X) + \epsilon$
Convex Optimization	We minimize $f_0(x)$. constraints- $f_i(x) \leq 0$, $1 \leq i \leq m$, $f_i(x) - convex$ $h_j(x) = 0$, $1 \leq j \leq l$, $h_j(x) - affine$ $L(X, \lambda, \nu) = f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^l \nu_i h_i(x)$. λ_i- Lagrange multiplier (for the inequalities) ν_i- Lagrange multiplier (for the equalities)
KKT conditions (conditions for $\tilde{x}$ to be a local minimum)	$\lambda_i - constants, 1 \leq i \leq m$ and $\nu_j, 1 \leq j \leq l$ $\nabla L(\tilde{x}, \lambda, \nu) = 0$ $f_i(\tilde{x}) \leq 0, 1 \leq i \leq m$ $h_j(\tilde{x}) = 0, 1 \leq j \leq l$ $\lambda_i f_i(\tilde{x}) = 0, 1 \leq i \leq m,$ $\lambda_i \geq 0$
Water filling	$P_i = [\frac{1}{\nu} - N_i]^+ = \max(0, \frac{1}{\nu} - N_i)$
Maximum entropy	$p(x) \geq 0, \forall x \in \mathcal{X}$. $\sum_{x \epsilon \mathcal{X}} p(x) = 1$. $\sum_{x \epsilon \mathcal{X}} p(x) r_i(x) = \alpha_i$, $i = (1, ..., m)$. $p(x) = e^{\lambda_0 - 1 + \sum_{i=1}^m \lambda_i r_i(x)}$
Cauchy-Schwarz	$E[X_1 X_2] \leq \sqrt{E[X_1^2]E[X_2^2]}$