Sylvester problem (prediction)

Assume: $X \subseteq Y$ convex, $D: Y \times Y \to \mathbb{R}_+$ satisfies compensation identity (e.g. $D = \|x-y\|^2$, $D = \sum p_i \ln \frac{p_i}{q_i}$)

Problem: Find y^* s.t. $R(y^*) = R_{min}$ $\left(= \inf_y R(y),\ R(y) = \sup_x D(x,y)\right)$.

Introduce: $\tilde{X} = \{\alpha = (\alpha_x)_{x\in X} \mid$ prob. dist. over X w. finite supp.$\}$,

$\alpha \rightsquigarrow \hat{\alpha}$ = barycenter $\left(\sum_x \alpha_x \cdot x\right)$, $\Phi(\alpha,y) = \sum \alpha_x D(x,y)$

Then: $\Phi(\alpha,y) = H(\alpha) + D(\alpha,y)$ with $H(\alpha) = \Phi(\alpha,\hat{\alpha}) = \sum \alpha_x D(x,\hat{\alpha})$
and $D(\alpha,y) = D(\hat{\alpha},y)$ and,

from theorem: $\boxed{H_{max} = R_{min}}$ (note: new $R(y) = \sup_\alpha \Phi(\alpha,y)$
$= \sup_\alpha \sum \alpha_x D(x,y) = \sup_x D(x,y) = R(y)$)

furthermore:

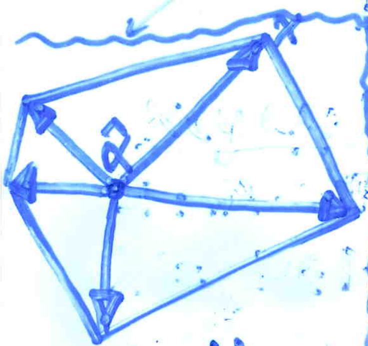

illustration of $H(\alpha)$, in information theory = information transition rate. Then H_{max} = capacity.

Kuhn-Tucker criterion. Given α^*, y^*, R such that $y^* = \widehat{\alpha^*}$, $D(x,y^*) \leq R$ for all $x \in X$, $D(x,y^*) = R$ for every **anchor** (i.e. an x with $\alpha_x > 0$). Then α^* and y^* are optimal.

(identification, 2nd case)

<u>Proof.</u> For any y, $R(y) = \sum_x \alpha^*_x \cdot R(y)$
$\geq \sum \alpha^*_x D(x,y) = \sum \alpha^*_x D(x,y^*) + D(y^*,y)$
$= R + D(y^*,y)$. Result follows as $R(y^*) = R$. $\blacksquare$