A Simple Multi-Class Boosting Framework – Supplement

Ron Appel, Pietro Perona

Claim 11: $\langle |\hat{\mathbf{v}}|, \mathbf{1} \rangle^2 \;\geq\; 1$

Proof: Reformulate as a constrained minimization problem, with $\mathbf{x} \in \mathbb{R}^N$:

$$\min_{\mathbf{x}} \{ \langle \mathbf{x}, \mathbf{1} \rangle \} \quad \text{such that: } \|\mathbf{x}\|^2 = 1, \; \mathbf{x} \geq \mathbf{0}$$

$$\therefore \; L = \langle \mathbf{x}, \mathbf{1} \rangle - \lambda(\|\mathbf{x}\|^2 - 1) - \sum_{n=1}^{N} \mu_n (\langle \mathbf{x}, \boldsymbol{\delta}_n \rangle - 0)$$

$$\text{such that: } \mu_n \geq 0 \;\; \forall n$$

$$\therefore \; \nabla_{\mathbf{x}} L = \mathbf{1} - 2\lambda\,\mathbf{x} - \sum_{n=1}^{N} \mu_n\, \boldsymbol{\delta}_n$$

$$\therefore \; 2\lambda\,\mathbf{x}^* = \sum_{n=1}^{N} (1-\mu_n)\,\boldsymbol{\delta}_n$$

$$\therefore \; \mathbf{x}^* \;=\; \frac{\sum_{n=1}^{N} (1-\mu_n)\,\boldsymbol{\delta}_n}{\sqrt{\sum_{n=1}^{N} (1-\mu_n)^2}} \qquad [\mu_n \leq 1 \; \forall\, n]$$

$$\therefore \; \langle \mathbf{x}^*, \mathbf{1} \rangle \;=\; \frac{\sum_{n=1}^{N} (1-\mu_n)}{\sqrt{\sum_{n=1}^{N} (1-\mu_n)^2}} \;\geq\; \frac{\sum_{n=1}^{N} (1-\mu_n)^2}{\sqrt{\sum_{n=1}^{N} (1-\mu_n)^2}} \;=\; \sqrt{\sum_{n=1}^{N} (1-\mu_n)^2}$$

To have unit norm, $\mathbf{x}$ must contain at least one non-zero element. Without loss of generality, we assume $x_1 > 0$; and hence: $\mu_1 = 0$

$$\therefore \; \langle \mathbf{x}^*, \mathbf{1} \rangle \;\geq\; \sqrt{1 + \sum_{n=2}^{N} (1-\mu_n)^2} \;\geq\; 1$$

Q.E.D.

Claim 12: $\max_i \{ \langle \mathbf{1} - 2\boldsymbol{\delta}_i, \hat{\mathbf{v}} \rangle^2 \} \;\geq\; \frac{4}{N}$ for $N \geq 4$

Proof: Reformulate as a constrained minimization problem with $\mathbf{x} \in \mathbb{R}^N$. Without loss of generality, assume that $\langle \mathbf{x}, \mathbf{1} \rangle \geq 0$ and that its first element x_1 is a minimal element (i.e. $x_1 \leq x_n \; \forall\, n$).

$$\min_{\mathbf{x}} \{ \langle \mathbf{x}, \mathbf{1} - 2\boldsymbol{\delta}_1 \rangle \} \quad \text{such that: } \|\mathbf{x}\|^2 = 1, \;\; \mathbf{x} \geq x_1\,\mathbf{1}$$

$$\therefore \; L = \langle \mathbf{x}, \mathbf{1} - 2\,\boldsymbol{\delta}_1 \rangle - \lambda(\|\mathbf{x}\|^2 - 1) - \sum_{n=2}^{N} \mu_n (\langle \mathbf{x}, \boldsymbol{\delta}_n \rangle - x_1)$$

$$\text{such that: } \mu_n \geq 0 \;\; \forall n$$

$$\therefore \; \nabla_{\mathbf{x}} L = [\mathbf{1} - 2\boldsymbol{\delta}_1] - 2\lambda\,\mathbf{x} - \sum_{n=2}^{N} \mu_n\, \boldsymbol{\delta}_n$$

$$\therefore \; 2\lambda\,\mathbf{x}^* = -\boldsymbol{\delta}_1 + \sum_{n=2}^{N} (1-\mu_n)\,\boldsymbol{\delta}_n$$

$$\therefore \; \mathbf{x}^* \;=\; \frac{-\boldsymbol{\delta}_1 + \sum_{n=2}^{N} (1-\mu_n)\,\boldsymbol{\delta}_n}{\sqrt{1 + \sum_{n=2}^{N} (1-\mu_n)^2}}$$

$$\therefore \; \langle \mathbf{x}^*, \mathbf{1} - 2\boldsymbol{\delta}_1 \rangle \;=\; \frac{1 + \sum_{n=2}^{N} (1-\mu_n)}{\sqrt{1 + \sum_{n=2}^{N} (1-\mu_n)^2}}$$

Note that if $x_n > x_1$ then $\mu_n = 0$, and if $x_n = x_1$ then $(1-\mu_n) = -1$. Let M be the number of unique indices $n \geq 2$ for which $x_n = x_1$.

$$\therefore \; \langle \mathbf{x}^*, \mathbf{1} - 2\boldsymbol{\delta}_1 \rangle \;=\; \frac{1 + \big((N-1) - M\big) - M}{\sqrt{N}} \;=\; \frac{N - 2M}{\sqrt{N}}$$

Since $\langle \mathbf{x}, \mathbf{1} \rangle \geq 0$
hence: $-1 + \big((N-1) - M\big) - M \geq 0 \;\; \therefore \; -2M \geq 2 - N$

$$\therefore \; \langle \mathbf{x}^*, \mathbf{1} - 2\boldsymbol{\delta}_1 \rangle \;\geq\; \frac{2}{\sqrt{N}} \qquad \therefore \; \langle \mathbf{x}^*, \mathbf{1} - 2\boldsymbol{\delta}_1 \rangle^2 \;\geq\; \frac{4}{N}$$

Q.E.D.