$$\text{F measure} := \frac{2\, Pr \ast Re}{Pr + Re}$$

$$Pr = \text{Precision} = \frac{\#\text{True positives}}{\#\text{True positives} + \#\text{False Positives}} = \frac{TP}{TP + FP}$$

$$Re = \text{Recall} := \frac{TP}{TP + FN}$$

$$1/F_\beta = \left(\beta/Pr + 1/Re\right) \frac{1}{2}$$

If $\beta > 1$ then you favour Recall

& o/w Precision

$$\min_{\mathbf{w}} \frac{1}{2}||\mathbf{w}||_2^2 + C \max_{\mathbf{y}'_1,\dots,\mathbf{y}'_N} \Big\{ \Delta\Big((\mathbf{y}_1,\dots,\mathbf{y}_N),(\mathbf{y}'_1,\dots,\mathbf{y}'_N)\Big)$$

$$+ \sum_{i=1}^{N} \max_{\mathbf{h}} \mathbf{w}\cdot\Phi(\mathbf{x}_i,\mathbf{h},\mathbf{y}'_i) - \sum_{i=1}^{N} \max_{\mathbf{h}} \mathbf{w}\cdot\Phi(\mathbf{x}_i,\mathbf{h},\mathbf{y}_i) \Big\}$$

$\min_{\omega} \frac{1}{2}$ [illegible]

convex : $g_1(w)$

$$-\sum_{i=1}^{N} \max_{h} \omega^T \phi(x_i, h_i, y_i)$$

concave : $-g_2(w)$

$g_1(w) - g_2(w)$ s.t g_1 & g_2 are both convex

Necessary conditions for optimality of $O(w) = g_1(w) - g_2(w)$

$$\nabla O(w^*) = \nabla g_1(w^*) - \nabla g_2(w^*) = 0$$

$$\nabla g_1(w^{t+1}) = \nabla g_2(w^t)$$ t = iteration number . . .

(set up our fix point iteration)

$$w^{t+1} = \min_w \left(g_1(w) - \underbrace{w^T \nabla g_2(w^t)}_{} \right)$$

More generally Majorization minimization techniques do the following . . .

$\min_{w} O(w)$ . . . . consider $T(w, w')$ s.t

$$O(w) \leq T(w, w') \quad \forall\, w'$$

$$\& \; O(w) = T(w, w)$$

w_0 .. initialized

$$w_{k+1} = \min_{w} T(w, w_k)$$

$$k = k+1$$

Majorization minimization algos . . .

Intuition behind convergence:

To prove: $O(w^{t+1}) \leq O(w^t)$

$$g_1(w^{t+1}) - g_2(w^{t+1}) \leq g_1(w^t) - g_2(w^t)$$

Rearranging

$$g_1(w^{t+1}) - g_1(w^t) \leq g_2(w^{t+1}) - g_2(w^t)$$

using

$$\nabla g_1(w^{t+1}) = \nabla g_2(w^t)$$

$$g_i(x') \geq g_i(x) + (x'-x)^T \nabla g_i(x)$$

for $i = 1$ & 2

For g_1, $x = w^{t+1}$ & $x' = w^t$

For g_2, $x = w^t$ & $x' = w^{t+1}$