Kelly Proof

Eugene Lykhovyd

1 Modern Kelly Algorithm Proof

Proof of the convergence (see [1]).

Theorem. Solving the problem

$$\min f(x) \tag{1}$$

subject to

$$x \in X$$

where $f(x)$ is continuous, convex function and X is a compact, convex set. Then the sequence $\{x_k\}$ of the solutions to

$$f_k(x) = \min \eta$$

subject to

$$\eta \geq \alpha_k(x), x \in X,$$

where $\alpha_k(x)$ - supporting hyperplane at x_{k-1}, converges to some $\bar{x}$, and, moreover, $f(\bar{x}) = f(x^*)$, where x^* is an optimal solution to the original problem (implies $\bar{x}$ is the optimal solution).

Proof.

It is only necessary to show, that there exists the subsequence of $f_k(x_k)$ which converges to f^{opt} - the optimum value of problem (1). If it does, then from monotonicity of $f_k(x_k)$ (because they are the solutions to the LP problem, and each next one has extra constraint added, so the optimum value is greater or equal to the previous one) it follows that $f_k(x_k)$ converges to the same limit.

The sequence $\{x_k\}$ is in X, and X is compact, so there exists the limit point of that sequence in X. In other words, there exists the subsequence $x_{k_i} \to \bar{x}$, where $\bar{x} \in X$ is the limit point.

For every x_{k_i}, $f_n(x_{k_i}) = f(x_{k_i}) \quad \forall n > k_i$, because the supporting hyperplane at point x_{k_i} was added at step k_i, and for any other point $x \in X$ the function $f(x)$ lies above all hyperplanes, so the equality continues to hold. This can be written as

$$f_k(x) \leq f(x) \quad \forall x \in X \tag{2}$$

- simply the supporting hyperplane inequality is true for all hyperplanes, thus true for the maximum of them. Since $f_k(x)$ might not converge to $f(x)$ even point-wise, we will use the simple trick to show that they are still equal at $\bar{x}$. We will construct the auxiliary function $\tilde{f}(x)$ in the following way. Note, that for every $x \in X$ the sequence $f_k(x)$ is bounded (for example by maximum of $f(x)$ on X - exists by Weierstrass theorem and $f_1(x_1)$). Also, it's monotonically non-decreasing. Thus, by monotone convergence theorem, there exist the limit, which we denote as $\tilde{f}(x)$. By construction, $f_k(x) \to \tilde{f}(x) \quad \forall x \in X$ - point-wise convergence. Since X is compact, we can apply Dini's theorem. So, $f_k \to \tilde{f}$ uniformly on X. Note, that f_k were piecewise linear functions, more precisely the maximum of supporting hyperplanes, and thus continuous functions. By uniform convergence theorem, the function $\tilde{f}$ is also continuous. Thus, $(x_{k_i} \to \bar{x}) \implies (\tilde{f}(x_{k_i}) \to \tilde{f}(\bar{x}))$.

From (2), when putting $k \to \infty$, we have that

$$\tilde{f}(x) \leq f(x) \quad \forall x \in X. \tag{3}$$

Also, let us show another convergence: $f_{k_i}(x_{k_i-1}) \to \tilde{f}(\bar{x})$. Fix $\epsilon \geq 0$. Firstly, $||\tilde{f}(x_{k_i-1}) - \tilde{f}(\bar{x})|| \leq \epsilon$, because $\tilde{f}$ is continuous. Secondly, $||f_{k_i-1}(x_{k_i}) - \tilde{f}(x_{k_i})|| \leq \epsilon$ because of uniform convergence (does not depend on x_{k_i} : $\sup|\tilde{f}(x) - f_{k_i-1}(x)| \leq \epsilon$). Then

$$||f_{k_i}(x_{k_i-1}) - \tilde{f}(\bar{x})|| \leq ||f_{k_i}(x_{k_i-1}) - \tilde{f}(x_{k_i-1})|| + ||\tilde{f}(x_{k_i-1}) - \tilde{f}(\bar{x})|| \leq 2\epsilon \tag{4}$$

Recall, that $f_{k_i}(x_{k_i-1}) = f(x_{k_i-1})$, from (4) $f(x_{k_i-1}) \to \tilde{f}(\bar{x})$. From continuous property of f we have that $f(x_k) \to f(\bar{x})$. As far as the limit is unique, we have that

$$f(\bar{x}) = \tilde{f}(\bar{x}) \tag{5}$$

Denote x^{opt} as the optimal solution for problem (1) (exists, because convex function on a compact). Then $f(x^{opt}) \leq f(\bar{x})$. From (2) we get

$$f_k(x^{opt}) \leq f(x^{opt}) \leq f(\bar{x}).$$

Recalling, that x_k is the minimum point for f_k, we have that $f_k(x_k) \leq f_k(x^{opt})$, so finally

$$f_{k_i}(x_{k_i}) \leq f(x^{opt}) \leq f(\bar{x})$$

Note that in (4) the x index is not that important because of uniform convergence, we have putting $k \to \infty$ and (5) that

$$\tilde{f}(\bar{x}) \leq f(x^{opt}) \leq \tilde{f}(\bar{x})$$

which implies that $f(\bar{x}) = f(x^{opt})$, and from this we get that $\bar{x}$ is the optimal solution.

2 References

[1] Kelley, J. E. J. "The cutting-plane method for solving convex programs." *Journal of the Society for Industrial and Applied Mathematics* (1960): 703-712.