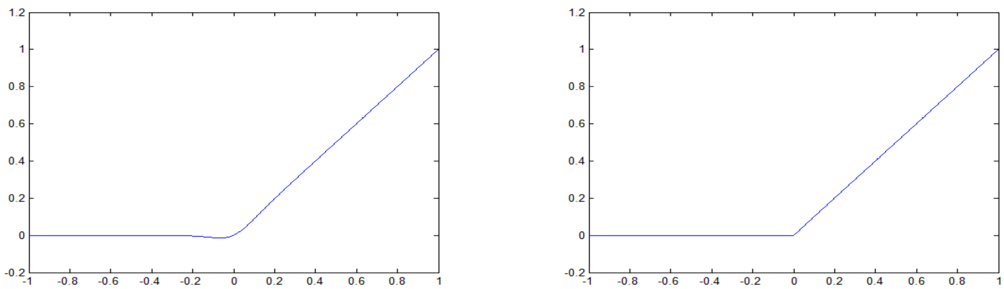
form to compute their Jacobian and Hessian. To overcome this problem, I approximate the nondifferentiable function $\max\{x, 0\}$ with a smooth function $\frac{x}{1+\exp(-x/h)}$ for small $h > 0$ and marginal distribution functions with finite normal mixtures $\sum_i a_i \Phi\left(\frac{x-\mu_i}{\sigma_i}\right)$, which makes it substantially simple to evaluate the Jacobian and Hessian of the objective function at any point.[26]



(a) $h = 0.05$ (b) $h = 0.01$

Figure B.2: Approximation of $\max\{x, 0\}$ and $\frac{x}{1+\exp(-x/h)}$

I used Knitro to solve the optimization problem using the smoothed functions. Knitro is a constrained nonlinear optimization software.[27] In optimization, I considered the constraints that $0 \leq a_{k+1} - a_k \leq \delta$ and $\delta < a_{k+2} - a_k$ for each integer k,and I fed the Jacobian and the Hessian of the Lagrangian into Knitro. Since the objective function in the optimization is not convex, it is likely to have multiple local maxima. I randomly generated initial values 90-200 times using the "multistart" feature in Knitro.

The numerical optimization results substantially depend on the initial values, which is the evidence of multiple local maxima and surprisingly, the values of the objective function at all these local maxima were lower than $V_K(\delta)$ in both Section 4 and Section 5. Based on the numerical evidence, it appears that the global maximum for both Section 4 and Section 5 is achieved or well approximated when $a_{k+1} - a_k = \delta$ for each integer k. It remains to show under which conditions on the joint distribution or marginal distributions the sharp lower bound is indeed achieved when $a_{k+1} - a_k = \delta$ for each integer k.

[26] I used the Kolmogorov-Smirnov test to determine the number of components in the mixture model. I increased the order of the mixture model from one until the test does not reject the null that the two distribution functions are identical. In the numerical example, I used one to three components for 9 different pairs of (k_1, k_2) considered in Section 4 and I used three for the empirical application. For each mixture model that I used to approximate the marginal distributions, the null hypothesis that two distribution functions are identical was not rejected with pvalue> 0.99.

[27] Recently Knitro has been often used to solve large-dimensional constrained optimization problems in the literature including Conlon (2012), Dubé et al. (2012) and Galichon and Salanié (2012). See Byrd et al. (2006) for details.