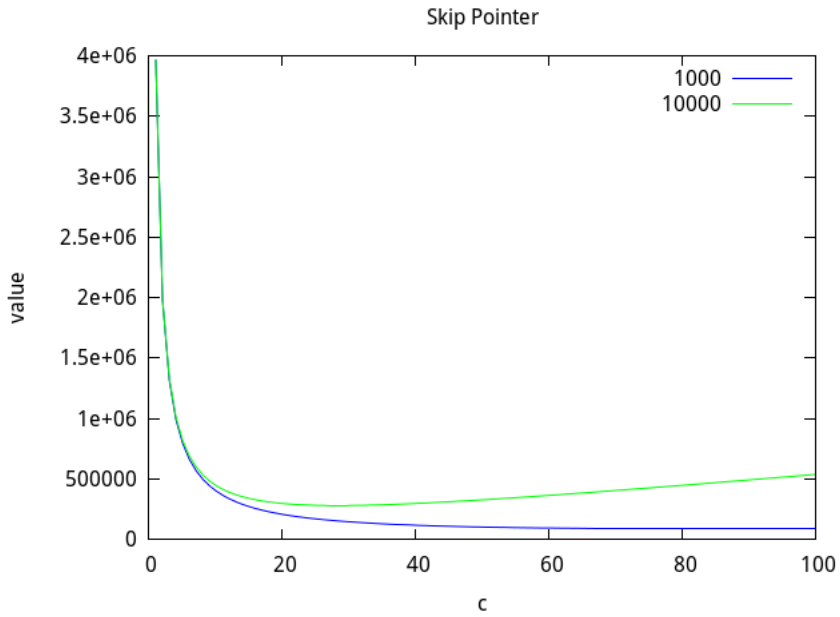
Information Retrieval: Homework #2

Lin Hung Cheng B01902059

Problem 1

Solution



$f(c) = \frac{kn}{c} + \frac{pc}{2}$
$f'(c) = \frac{p}{2} - \frac{kn}{c^2}$
when minimized value, $f'(c) = 0$, $c = \frac{\sqrt{2kn}}{\sqrt{p}}$

Problem 2

Exercise 6.12
How does the base of the logarithm in (6.7) affect the score calculation in (6.9)? How does the base of the logarithm affect the relative scores of two documents on a given query?
Solution
the precise base of the logarithm has not effect to ranking, but rescale the score.
$idf(b) = \log_b(\frac{N}{df}) = \frac{\log_{10}\frac{N}{df}}{\log_{10} b}$
$\frac{idf(b)}{idf(10)} = \frac{1}{b\log_{10}}$ is a constant
So score will be divided by $\log_{10}(b)$, and the query result will not change