Analytical Form of Gradient and Hessian

This document illustrates the analytical form of gradient and hessian for the forward intensity model. For brevity, we only deal with the case of default. Gradient and hessian for the intensity function corresponding to other exit can be derived similarly.

For notational convenience, we introduce the following notations:
β denotes the parameter vector (column vector)
We classify all the firm-month observations into three categories:

$$X^0 = \begin{bmatrix} x_1^0 \\ x_2^0 \\ \cdots \\ x_{N_0}^0 \end{bmatrix}, X^1 = \begin{bmatrix} x_1^1 \\ x_2^1 \\ \cdots \\ x_{N_1}^1 \end{bmatrix}, X^2 = \begin{bmatrix} x_1^2 \\ x_2^2 \\ \cdots \\ x_{N_2}^2 \end{bmatrix}$$

X^0 contains all the surviving firm-month observations, X^1 contains all the default observations and X^2 has all the observations for other exit. There are altogether N_0, N_1 and N_2 observations for each category. Variables for each firm-month observation forms a row vector x_i^j and we assume the first variable is always a constant 1 corresponding to the intercept term in the intensity function.

The case without bailout term

The pseudo log-likelihood function is expressed as follows:

$$L = -\sum_{i=1}^{N_0} \exp(x_i^0\beta)\Delta t + \sum_{i=1}^{N_1} \log\left\{1 - \exp\left[-\exp\left(x_i^1\beta\right)\Delta t\right]\right\} - \sum_{i=1}^{N_2} \exp(x_i^2\beta)\Delta t$$

Let $f_i^j = \exp(x_i^j\beta)\Delta t$, then we can express gradient and hessian as follows:

$$\text{gradient} = -\sum_{i=1}^{N_0} f_i^0 (x_i^0)^T + \sum_{i=1}^{N_1} \frac{\exp(-f_i^1) f_i^1}{1-\exp(-f_i^1)} (x_i^1)^T - \sum_{i=1}^{N_2} f_i^2 (x_i^2)^T$$

$$\text{hessian} = -\sum_{i=1}^{N_0} f_i^0 (x_i^0)^T (x_i^0) + \sum_{i=1}^{N_1} \frac{\exp(-f_i^1) f_i^1 [1 - f_i^1 - \exp(-f_i^1)]}{[1-\exp(-f_i^1)]^2} (x_i^1)^T (x_i^1) - \sum_{i=1}^{N_2} f_i^2 (x_i^2)^T (x_i^2)$$

where $(x_i^j)^T$ denotes the transpose of x_i^j

The case with bailout term

In this case, we assume the bailout effect starts to take place after t_{bailout} and t_i^j denotes the time of the observation x_i^j. We also assume the first two parameters $\beta(1)$ and $\beta(2)$ are the bailout parameters. With slight difference from the paper, we assume the intensity function to have the following form:

$$\text{intensity}_i^j = \exp\{-\beta(1)^2 \exp[-\beta(2)^2 (t_i^j - t_{\text{bailout}})] 1_{t_i^j > t_{\text{bailout}}} + x_i^j \beta(3:\text{end})\}$$

To employ $-\beta(1)^2$ and $\beta(2)^2$, we can guarantee the two bailout parameters to be negative and positive respectively so that they are consistent with their definition in the paper. To make the presentation easier, we further introduce the following notations:

$$f_i^j = \exp\{-\beta(1)^2 \exp[-\beta(2)^2(t_i^j - t_{\text{bailout}})]1_{t_i^j > t_{\text{bailout}}} + x_i^j\beta(3:\text{end})\}\Delta t$$

$$g_i^j = \beta(1)^2 \exp[-\beta(2)^2(t_i^j - t_{\text{bailout}})]1_{t_i^j > t_{\text{bailout}}}$$

$$h_i^j = 2\beta(1) \exp[-\beta(2)^2(t_i^j - t_{\text{bailout}})]1_{t_i^j > t_{\text{bailout}}}$$

$$k_i^j = 2 \exp[-\beta(2)^2(t_i^j - t_{\text{bailout}})]1_{t_i^j > t_{\text{bailout}}}$$

$$m_i^j = 2\beta(2)(t_i^j - t_{\text{bailout}})$$

$$n_i^j = 2(t_i^j - t_{\text{bailout}})$$

In this case, gradient(3:end) and hessian(3:end,3:end) can be derived similarly as in the case without bailout term:

$$\text{gradient}(3:\text{end}) = -\sum_{i=1}^{N_0} f_i^0 (x_i^0)^T + \sum_{i=1}^{N_1} \frac{\exp(-f_i^1) f_i^1}{1 - \exp(-f_i^1)} (x_i^1)^T - \sum_{i=1}^{N_2} f_i^2 (x_i^2)^T$$

$$\begin{aligned}\text{hessian}(3:\text{end}, 3:\text{end}) = &-\sum_{i=1}^{N_0} f_i^0 (x_i^0)^T (x_i^0) + \sum_{i=1}^{N_1} \frac{\exp(-f_i^1) f_i^1 [1 - f_i^1 - \exp(-f_i^1)]}{[1 - \exp(-f_i^1)]^2} (x_i^1)^T (x_i^1) \\ &- \sum_{i=1}^{N_2} f_i^2 (x_i^2)^T (x_i^2)\end{aligned}$$

We now derive other elements in the gradient and hessian:

$$\text{gradient}(1) = \sum_{i=1}^{N_0} f_i^0 h_i^0 - \sum_{i=1}^{N_1} \frac{\exp(-f_i^1) f_i^1 h_i^1}{1 - \exp(-f_i^1)} + \sum_{i=1}^{N_2} f_i^2 h_i^2$$

$$\text{gradient}(2) = -\sum_{i=1}^{N_0} f_i^0 g_i^0 m_i^0 + \sum_{i=1}^{N_1} \frac{\exp(-f_i^1) f_i^1 g_i^1 m_i^1}{1 - \exp(-f_i^1)} - \sum_{i=1}^{N_2} f_i^2 g_i^2 m_i^2$$

$$\begin{aligned}\text{hessian}(1,1) = &\sum_{i=1}^{N_0} f_i^0 [k_i^0 - (h_i^0)^2] \\ &- \sum_{i=1}^{N_1} \frac{\exp(-f_i^1)(f_i^1)^2(h_i^1)^2 + [1 - \exp(-f_i^1)] \exp(-f_i^1) f_i^1 [k_i^1 - (h_i^1)^2]}{[1 - \exp(-f_i^1)]^2} \\ &+ \sum_{i=1}^{N_2} f_i^2 [k_i^2 - (h_i^2)^2]\end{aligned}$$

$$\begin{aligned}\text{hessian}(2,2) = &\sum_{i=1}^{N_0} f_i^0 g_i^0 [(m_i^0)^2 (1 - g_i^0) - n_i^0] \\ &- \sum_{i=1}^{N_1} \frac{\exp(-f_i^1)(f_i^1)^2(g_i^1)^2(m_i^1)^2 - [1 - exp(-f_i^1)] \exp(-f_i^1) f_i^1 g_i^1 [(g_i^1 - 1)(m_i^1)^2 + n_i^1]}{[1 - \exp(-f_i^1)]^2} \\ &+ \sum_{i=1}^{N_2} f_i^2 g_i^2 [(m_i^2)^2 (1 - g_i^2) - n_i^2]\end{aligned}$$

$$
\begin{aligned}
\text{hessian}(1,2) = &\sum_{i=1}^{N_0} f_i^0 h_i^0 m_i^0 (g_i^0 - 1) \\
&+ \sum_{i=1}^{N_1} \frac{\exp(-f_i^1)(f_i^1)^2 g_i^1 h_i^1 m_i^1 - [1-\exp(-f_i^1)]\exp(-f_i^1) f_i^1 h_i^1 m_i^1 (g_i^1 - 1)}{[1-\exp(-f_i^1)]^2} \\
&+ \sum_{i=1}^{N_2} f_i^2 h_i^2 m_i^2 (g_i^2 - 1)
\end{aligned}
$$

$$
\text{hessian}(1,3:\text{end}) = \sum_{i=1}^{N_0} f_i^0 h_i^0 x_i^0 - \sum_{i=1}^{N_1} \frac{\exp(-f_i^1) f_i^1 [1 - f_i^1 - \exp(-f_i^1)] h_i^1}{[1-\exp(-f_i^1)]^2} x_i^1 + \sum_{i=1}^{N_2} f_i^2 h_i^2 x_i^2
$$

$$
\begin{aligned}
\text{hessian}(2,3:\text{end}) = &-\sum_{i=1}^{N_0} f_i^0 g_i^0 m_i^0 x_i^0 + \sum_{i=1}^{N_1} \frac{\exp(-f_i^1) f_i^1 [1 - f_i^1 - \exp(-f_i^1)] g_i^1 m_i^1}{[1-\exp(-f_i^1)]^2} x_i^1 \\
&- \sum_{i=1}^{N_2} f_i^2 g_i^2 m_i^2 x_i^2
\end{aligned}
$$