Функция распределения

1. По определению

$$F(x)=\int\limits_{-\infty}^{x}\frac{1}{\sum\limits_{i=0}^{k}\sum\limits_{j=0}^{k}\alpha_i\alpha_j\mathsf{E}X^{i+j}}\frac{e^{-\frac{t-\alpha}{\beta}}}{\beta(1+e^{-\frac{t-\alpha}{\beta}})^2}(\sum_{i=0}^{k}\alpha_i t^i)^2dt=$$

$$=\frac{1}{\beta\sum\limits_{i=0}^{k}\sum\limits_{j=0}^{k}\alpha_i\alpha_j\mathsf{E}X^{i+j}}\int\limits_{-\infty}^{x}\sum_{i=0}^{k}\sum_{j=0}^{k}\alpha_i\alpha_j t^{i+j}\frac{e^{-\frac{t-\alpha}{\beta}}}{(1+e^{-\frac{t-\alpha}{\beta}})^2}dt=$$

$$=\mathsf{C_1}\sum_{i=0}^{k}\sum_{j=0}^{k}\alpha_i\alpha_j\int\limits_{-\infty}^{x}\frac{1}{\beta}t^{i+j}\frac{e^{-\frac{t-\alpha}{\beta}}}{(1+e^{-\frac{t-\alpha}{\beta}})^2}dt=$$

$$=\mathsf{C_1}\begin{pmatrix}\alpha_0 & \alpha_1 & \dots & \alpha_k\end{pmatrix}\begin{pmatrix}\alpha_0 & \alpha_1 & \dots & \alpha_k & 0 & \dots & 0\\ 0 & \alpha_0 & \alpha_1 & \dots & \alpha_k & \dots & 0\\ \vdots & & \vdots & & \vdots & \ddots & \vdots\\ 0 & \dots & 0 & \alpha_0 & \alpha_1 & \dots & \alpha_k\end{pmatrix}\begin{pmatrix}\mathsf{I}^{\mathsf{t}}_0\\ \mathsf{I}^{\mathsf{t}}_1\\ \vdots\\ \mathsf{I^t}_{2k}\end{pmatrix}$$

2. Рассмотрим в общем виде интеграл. l определяется параметрами аппроксимации.

$$\mathsf{I^t}_l=\int\limits_{-\infty}^{x}\frac{1}{\beta}t^l\frac{e^{-\frac{t-\alpha}{\beta}}}{(1+e^{-\frac{t-\alpha}{\beta}})^2}dt=$$

Введем замену: $\frac{t-\alpha}{\beta}=u\Longrightarrow t=\beta u+\alpha,\ y=\frac{x-\alpha}{\beta}$

$$=\int\limits_{-\infty}^{y}\frac{1}{\beta}\beta(\beta u+\alpha)^l\frac{e^{-u}}{(1+e^{-u})^2}du=\int\limits_{-\infty}^{y}\sum_{i=0}^{l}C_l^i\alpha^{l-i}\beta^iu^i\frac{e^{-u}}{(1+e^{-u})^2}du=$$

$$=\sum_{i=0}^{l}C_l^i\alpha^{l-i}\beta^i\int\limits_{-\infty}^{y}u^i\frac{e^{-u}}{(1+e^{-u})^2}du=$$

$$=\begin{pmatrix}C_l^0\alpha^l & C_l^1\alpha^{l-1}\beta^1 & \dots & C_l^l\beta^l & 0 & \dots & 0\end{pmatrix}\begin{pmatrix}\mathsf{I}^{\mathsf{u}}_0\\ \mathsf{I}^{\mathsf{u}}_1\\ \vdots\\ \mathsf{I^u}_{2k}\end{pmatrix}$$

Тогда

$$\begin{pmatrix}\mathsf{I}^{\mathsf{t}}_0\\ \mathsf{I}^{\mathsf{t}}_1\\ \vdots\\ \mathsf{I^t}_{2k}\end{pmatrix}=\begin{pmatrix}C_0^0\alpha^0\beta^0 & 0 & \dots & 0\\ C_1^0\alpha\beta^0 & C_1^1\alpha^0\beta & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ C_{2k}^0\alpha^{2k}\beta^0 & C_{2k}^1\alpha^{2k-1}\beta & \dots & C_{2k}^{2k}\alpha^0\beta^{2k}\end{pmatrix}\begin{pmatrix}\mathsf{I}^{\mathsf{u}}_0\\ \mathsf{I}^{\mathsf{u}}_1\\ \vdots\\ \mathsf{I^u}_{2k}\end{pmatrix}$$

3. Очередной интеграл, его нахождение и есть основная часть работы. До гиперболического секанса доходим по определению момента логистического распределения как в Вики:

$$\mathsf{I^u}_m = \int\limits_{-\infty}^{y} u^m \tfrac{e^{-u}}{(1+e^{-u})^2} du =$$

$$= \int\limits_{-\infty}^{y} u^m \tfrac{e^u e^{-u}}{e^u(1+2e^{-u}+e^{-2u})} du = \int\limits_{-\infty}^{y} u^m \tfrac{1}{(e^u+2+e^{-u})} du = \int\limits_{-\infty}^{y} u^m \tfrac{1}{4} \tfrac{4}{(e^{\frac{u}{2}}+e^{-\frac{u}{2}})^2} du =$$

$$= \tfrac{1}{4} \int\limits_{-\infty}^{y} u^m \operatorname{sech}^2 \tfrac{u}{2} du$$

Функция симметрична относительно начала координат, если m нечетно, симметрична относительно оси Oy, если m четно. Тут уже

$$\mathsf{I^u}_m = (-1)^m \tfrac{1}{4} \int\limits_{0}^{\infty} u^m \operatorname{sech}^2 \tfrac{u}{2} du + (-1)^{[m \equiv 0 \pmod 2][y<0]} \tfrac{1}{4} \int\limits_{0}^{|y|} u^m \operatorname{sech}^2 \tfrac{u}{2} du =$$

Первое слагаемое – считается легко, зависит от только от порядка (порядком считаю степень переменной, которая определяется параметрами аппроксимации). Нашла в статье (1), ссылку прикреплю позже. Второе слагаемое меняется в зависимости от аргумента функции распределения.

Обозначим $\mathsf{I^v}_m = \int\limits_{0}^{v} u^m \operatorname{sech}^2 \tfrac{u}{2} du$,

$$= (-1)^m \tfrac{1}{4} \mathsf{I^{v=inf}}_m + (-1)^{[m \equiv 0 \pmod 2][y<0]} \tfrac{1}{4} \mathsf{I^{v=|y|}}_m$$

Рассмотрим два частных случая (это действительно важно, тут есть слагаемые, которые с ростом порядка исчезают, общую формулу вывести трудно, только если индикаторы добавить):

4. Пусть m = 0

$$\mathsf{I^v}_0 = \int\limits_{0}^{v} \operatorname{sech}^2 \tfrac{u}{2} du = 2 \int\limits_{0}^{v} d \tanh \tfrac{u}{2} = 2 \tanh \tfrac{v}{2}$$

5. Пусть m = 1

$$\mathsf{I^v}_1 = \int\limits_{0}^{v} u \operatorname{sech}^2 \tfrac{u}{2} du =$$

$$= 2 \int\limits_{0}^{v} u d \tanh \tfrac{u}{2} = 2v \tanh \tfrac{v}{2} - 4 \int\limits_{0}^{v} \tfrac{1}{2} \tanh \tfrac{v}{2} du = 2v \tanh \tfrac{v}{2} - 4 \ln \cosh \tfrac{v}{2} =$$

$$= 2v \tanh \tfrac{v}{2} - 2v + 4 \operatorname{Li}_1(-e^{-v}) + 4 \ln 2$$

Дальше до пункта 7 нет шагов сложнее интегралов по частям и использования определения и базовых свойств полилогарифма, которые легко гуглятся и есть в вольфраме, добавлю позже.

6. Пусть $m \geq 2$

$$\mathsf{I^v}_m = \int\limits_{0}^{v} u^m \operatorname{sech}^2 \tfrac{u}{2} du =$$

$$= 2\int\limits_0^v u^m d\tanh\tfrac{u}{2} = 2v^m\tanh\tfrac{v}{2} - 2\int\limits_0^v mu^{m-1}\tanh\tfrac{u}{2}du =$$

$$= 2v^m\tanh\tfrac{v}{2} - 4m\int\limits_0^v u^{m-1}d\ln\cosh\tfrac{u}{2} =$$

$$= 2v^m\tanh\tfrac{v}{2} - 4mv^{m-1}\ln\cosh\tfrac{v}{2} + 4m\int\limits_0^v (m-1)u^{m-2}\ln\cosh\tfrac{u}{2}du =$$

$$= 2v^m\tanh\tfrac{v}{2} - 4mv^{m-1}(\tfrac{v}{2} - \mathrm{Li}_1(-e^{-v}) - \ln 2) + 4m(m-1)\int\limits_0^v u^{m-2}(\tfrac{u}{2} - \mathrm{Li}_1(-e^{-u}) - \ln 2)du =$$

$$= 2v^m\tanh\tfrac{v}{2} - 2v^m + 4mv^{m-1}\mathrm{Li}_1(-e^{-v}) - 4m(m-1)\int\limits_0^v u^{m-2}\mathrm{Li}_1(-e^{-u})du =$$

$$= 2v^m\tanh\tfrac{v}{2} - 2v^m + 4mv^{m-1}\mathrm{Li}_1(-e^{-v}) - 4m(m-1)\int\limits_0^v \mathrm{Li}_1(-e^{-u})d\tfrac{u^{m-1}}{m-1} =$$

$$= 2v^m\tanh\tfrac{v}{2} - 2v^m + 4mv^{m-1}\mathrm{Li}_1(-e^{-v}) - 4mu^{m-1}\mathrm{Li}_1(-e^{-u}) - 4m(m-1)\int\limits_0^v \tfrac{u^{m-1}}{m-1}d\ln(1+e^{-u}) =$$

$$= 2v^m\tanh\tfrac{v}{2} - 2v^m + 4m\int\limits_0^v u^{m-1}\tfrac{1}{1+e^u}du =$$

Обозначим $\mathsf{I^f}_m = \int\limits_0^v u^m\tfrac{1}{1+e^u}du$

$$= 2v^m\tanh\tfrac{v}{2} - 2v^m + 4m\mathsf{I^f}_{m-1}$$

7. Самое веселое начинается здесь. Переход между второй и третьей строкой надо пояснить формулой из статьи (2).

$$\mathsf{I^f}_m = \int\limits_0^v u^m\tfrac{1}{1+e^u}du =$$

$$= \int\limits_0^\infty u^m\tfrac{1}{1+e^u}du - \int\limits_v^\infty u^m\tfrac{1}{1+e^u}du =$$

$$= \Gamma(m+1)(1-2^{-m})\zeta(m+1) - \int\limits_v^\infty u^m\tfrac{1}{1+e^u}du =$$

$$= \Gamma(m+1)(1-2^{-m})\zeta(m+1) - \int\limits_0^\infty (v+u)^m\tfrac{1}{1+e^{v+u}}du =$$

$$= \Gamma(m+1)(1-2^{-m})\zeta(m+1) - \sum_{i=0}^m C_m^i v^{m-i}\int\limits_0^\infty u^i\tfrac{1}{1+e^{v+u}}du =$$

Заметим, что $\frac{1}{1+e^u} = \sum\limits_{j=1}^\infty (-1)^{j-1}e^{-ju}$

$$= \Gamma(m+1)(1-2^{-m})\zeta(m+1) - \sum_{i=0}^m C_m^i v^{m-i}\int\limits_0^\infty u^i\sum_{j=1}^\infty(-1)^{j-1}e^{-j(v+u)}du =$$

$$= \Gamma(m+1)(1-2^{-m})\zeta(m+1) - \sum_{i=0}^{m} C_m^i v^{m-i} \sum_{j=1}^{\infty} (-1)^{j-1} e^{-jv} \int_0^{\infty} u^i e^{-ju} du =$$

$$= \Gamma(m+1)(1-2^{-m})\zeta(m+1) - \sum_{i=0}^{m} C_m^i v^{m-i} \sum_{j=1}^{\infty} (-1)(-1)^j (e^{-v})^j \int_0^{\infty} \tfrac{1}{j^{i+1}} (ju)^i e^{-ju} d(ju) =$$

Поскольку $\int_0^{\infty} u^i e^{-u} du = \Gamma(i+1)$

$$= \Gamma(m+1)(1-2^{-m})\zeta(m+1) + \sum_{i=0}^{m} \tfrac{\Gamma(m+1)}{\Gamma(i+1)\Gamma(m-i+1)} v^{m-i} \sum_{j=1}^{\infty} \tfrac{(-e^{-v})^j}{j^{i+1}} \Gamma(i+1) =$$

По определению $\mathrm{Li_p(z)} = \sum_{j=1}^{\infty} \frac{z^j}{j^p}$

$$= \Gamma(m+1)(1-2^{-m})\zeta(m+1) + \Gamma(m+1) \sum_{i=0}^{m} \tfrac{v^{m-i}}{\Gamma(m-i+1)} \mathrm{Li_{i+1}}(-e^{-v})$$

Найдем несколько констант.

8. Если $v \longrightarrow \infty$, то

$$\mathsf{I^v}_m = \int_0^v u^m \operatorname{sech}^2 \tfrac{u}{2} du$$

$$\lim_{v\to\infty} \tfrac{1}{4} \mathsf{I^v}_0 = \lim_{v\to\infty} \tfrac{1}{2} \tanh \tfrac{v}{2} = \tfrac{1}{2}$$

$$\lim_{v\to\infty} \tfrac{1}{4} \mathsf{I^v}_1 = \lim_{v\to\infty} (\tfrac{1}{2} v \tanh \tfrac{v}{2} - \tfrac{1}{2} v + \mathrm{Li}_1(-e^{-v}) + \ln 2) = \ln 2$$

$$\lim_{v\to\infty} \tfrac{1}{4} \mathsf{I^v}_m = \lim_{v\to\infty} (\tfrac{1}{2} v^m \tanh \tfrac{v}{2} - \tfrac{1}{2} v^m) + m\Gamma(m)(1-2^{1-m})\zeta(m) = m\Gamma(m)(1-2^{1-m})\zeta(m)$$

9. Сведем все к матрицам. Обозначим $\mathsf{C_m} = m\Gamma(m)(1-2^{1-m})\zeta(m)$

$$\mathsf{I^u}_m = (-1)^m \mathsf{C_m} +$$
$$(-1)^{[m\equiv 0 \pmod 2][y<0]} \left(\tfrac{1}{2}|y|^m \tanh \tfrac{|y|}{2} - \tfrac{1}{2}|y|^m + \mathsf{C_m} + m\Gamma(m) \sum_{i=0}^{m-1} \tfrac{|y|^{m-i-1}}{\Gamma(m-i)} \mathrm{Li_{i+1}}(-e^{-|y|}) \right) =$$

$$= \mathsf{C_m} \left((-1)^m + (-1)^{[m\equiv 0 \pmod 2][y<0]} \right) +$$
$$(-1)^{[m\equiv 0 \pmod 2][y<0]} \left(\tfrac{1}{2}|y|^m (\tanh \tfrac{|y|}{2} - 1) + m\Gamma(m) \sum_{i=0}^{m-1} \tfrac{|y|^{m-i-1}}{\Gamma(m-i)} \mathrm{Li_{i+1}}(-e^{-|y|}) \right) =$$

Обозначим $\mathsf{C^y_m} = (-1)^{[m\equiv 0 \pmod 2][y<0]}$, $\mathsf{C^{inf}_m} = (-1)^m$

$$= \mathsf{C_m}(\mathsf{C^{inf}_m} + \mathsf{C^y_m}) + \mathsf{C^y_m} \left(\tfrac{1}{2}|y|^m (\tanh \tfrac{|y|}{2} - 1) + m\Gamma(m) \sum_{i=0}^{m-1} \tfrac{|y|^{m-i-1}}{\Gamma(m-i)} \mathrm{Li_{i+1}}(-e^{-|y|}) \right)$$

$$\mathsf{I^u}_0 = \lim_{y\to\infty} \tfrac{1}{2}|y|^m \tanh \tfrac{|y|}{2} + \mathsf{C^y_m} \tfrac{1}{2}|y|^m \tanh \tfrac{|y|}{2}$$

$$\begin{pmatrix} \mathsf{I}^{\mathsf{u}}_1 \\ \mathsf{I}^{\mathsf{u}}_2 \\ \vdots \\ \mathsf{I^u}_{2k} \end{pmatrix} = \mathsf{C} \circ (\mathsf{C}^{\mathsf{inf}} + \mathsf{C}^{\mathsf{y}}) + \mathsf{C}^{\mathsf{y}} \circ \left(\frac{1}{2}\tanh\frac{|y|}{2} - \frac{1}{2}\right) \begin{pmatrix} |y|^1 \\ |y|^2 \\ \vdots \\ |y|^{2k} \end{pmatrix} + \mathsf{C}^{\mathsf{y}} \circ \begin{pmatrix} \Gamma(1) \\ 2\Gamma(2) \\ \vdots \\ 2k\Gamma(2k) \end{pmatrix} \circ \mathsf{C}^{\mathsf{Li}}$$

Где $\circ$ – символ поэлементного умножения,

$$\mathsf{C}^{\mathsf{Li}} = \begin{pmatrix} \frac{1}{\Gamma(1)} & 0 & \dots & 0 \\ \frac{|y|}{\Gamma(2)} & \frac{1}{\Gamma(1)} & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ \frac{|y|^{2k-1}}{\Gamma(2k)} & \frac{|y|^{2k-2}}{\Gamma(2k-1)} & \dots & \frac{1}{\Gamma(1)} \end{pmatrix} \begin{pmatrix} \mathrm{Li}_1(-e^{-|y|}) \\ \mathrm{Li}_2(-e^{-|y|}) \\ \vdots \\ \mathrm{Li}_{2\mathrm{k}}(-e^{-|y|}) \end{pmatrix}$$

Побочное: посчитаются моменты, очев лучше, чем в сайпае. Для оптимизации понадобятся градиенты, все уже найдено, градиенты перепроверить численно.

10. Моменты

$$\sum_{i=0}^{k}\sum_{j=0}^{k} \alpha_i\alpha_j \mathsf{E}X^{i+j} =$$

$$= \begin{pmatrix} \alpha_0 & \alpha_1 & \dots & \alpha_k \end{pmatrix} \begin{pmatrix} \alpha_0 & \alpha_1 & \dots & \alpha_k & 0 & \dots & 0 \\ 0 & \alpha_0 & \alpha_1 & \dots & \alpha_k & \dots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \alpha_0 & \alpha_1 & \dots & \alpha_k \end{pmatrix} \begin{pmatrix} \mathsf{E}X^0 \\ \mathsf{E}X^1 \\ \vdots \\ \mathsf{E}X^{2k} \end{pmatrix}$$

$$\mathsf{E}X^q = \int\limits_{-\infty}^{\infty} x^q \frac{e^{-\frac{x-\alpha}{\beta}}}{\beta(1+e^{-\frac{x-\alpha}{\beta}})^2} dx$$

Введем замену: $\frac{x-\alpha}{\beta} = y \Longrightarrow x = \beta y + \alpha$

$$= \int\limits_{-\infty}^{\infty} \tfrac{1}{\beta}\beta(\beta y + \alpha)^q \tfrac{e^{-y}}{(1+e^{-y})^2} dy = \int\limits_{-\infty}^{\infty} \sum_{i=0}^{q} C_q^i \alpha^{q-i}\beta^i y^i \tfrac{e^{-y}}{(1+e^{-y})^2} dy =$$

$$= \sum_{i=0}^{q} C_q^i \alpha^{q-i}\beta^i \int\limits_{-\infty}^{\infty} y^i \tfrac{e^{-y}}{(1+e^{-y})^2} dy =$$

$$= \begin{pmatrix} C_q^0\alpha^q & C_q^1\alpha^{q-1}\beta^1 & \dots & C_q^q\beta^q & 0 & \dots & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{2} + (-1)^0\frac{1}{2} \\ \ln 2 + (-1)^1 \ln 2 \\ \vdots \\ 2k\Gamma(2k)(1-2^{2k})\zeta(2k) + (-1)^{2k} 2k\Gamma(2k)(1-2^{2k})\zeta(2k) \end{pmatrix}$$

11. Градиент

$$L(\alpha, w; x, y) = -\sum_{i=0}^{l} y_i \ln F(\alpha, \langle w, x_i\rangle) + (1-y_i)\ln\left(1 - F(\alpha, \langle w, x_i\rangle)\right)$$

Пусть $\theta = (\alpha, w)$, обозначим $\mathsf{F_i} = F(\alpha, \langle w, x_i \rangle)$

$$\frac{\partial}{\partial\theta}L(\alpha, w; x, y) = -\sum_{i=0}^{l}\frac{\partial}{\partial\theta}\left(y_i \ln \mathsf{F_i} + (1-y_i)\ln(1-\mathsf{F_i})\right) =$$

$$= \sum_{i=0}^{l}\frac{-y_i}{\mathsf{F_i}}\frac{\partial}{\partial\theta}\mathsf{F_i} + \frac{1-y_i}{1-\mathsf{F_i}}\frac{\partial}{\partial\theta}\mathsf{F_i} = \sum_{i=0}^{l}\frac{-y_i+y_i\mathsf{F_i}+\mathsf{F_i}-y_i\mathsf{F_i}}{\mathsf{F_i}(1-\mathsf{F_i})}\frac{\partial}{\partial\theta}\mathsf{F_i} =$$

$$= \sum_{i=0}^{l}\frac{\mathsf{F_i}-y_i}{\mathsf{F_i}(1-\mathsf{F_i})}\frac{\partial}{\partial\theta}\mathsf{F_i} = \sum_{i=0}^{l}\frac{\mathsf{F_i}-y_i}{\mathsf{F_i}(1-\mathsf{F_i})}\left(\frac{\partial}{\partial\alpha}\mathsf{F_i}\,, \frac{\partial}{\partial w}\mathsf{F_i}\right)$$

12.

$$\frac{\partial}{\partial\alpha}\mathsf{F_i} = \frac{\partial}{\partial\alpha}\frac{1}{\sum\limits_{i=0}^{k}\sum\limits_{j=0}^{k}\alpha_i\alpha_j \mathsf{E}X^{i+j}}\sum_{i=0}^{k}\sum_{j=0}^{k}\alpha_i\alpha_j\int\limits_{-\infty}^{\langle w\,,x_{\mathsf{i}}\rangle}\frac{1}{\beta}t^{i+j}\frac{e^{-\frac{t-\alpha}{\beta}}}{(1+e^{-\frac{t-\alpha}{\beta}})^2}dt =$$

$$= \frac{\partial}{\partial\alpha}\frac{\sum\limits_{i=0}^{k}\sum\limits_{j=0}^{k}\alpha_i\alpha_j \mathsf{I^{ti}}_{i+j}}{\sum\limits_{i=0}^{k}\sum\limits_{j=0}^{k}\alpha_i\alpha_j \mathsf{E}X^{i+j}}$$

$$\frac{\partial}{\partial\alpha}\sum_{i=0}^{k}\sum_{j=0}^{k}\alpha_i\alpha_j\mathsf{I^t}_{i+j} = (2\sum_{j=0}^{k}\alpha_j\mathsf{I^t}_{i+j})_{i=0}^{k} = 2\begin{pmatrix}\alpha_0 & \alpha_1 & \dots & \alpha_k & 0 & \dots & 0\\ 0 & \alpha_0 & \alpha_1 & \dots & \alpha_k & \dots & 0\\ \vdots & & \vdots & & \vdots & \ddots & \vdots\\ 0 & \dots & 0 & \alpha_0 & \alpha_1 & \dots & \alpha_k\end{pmatrix}\begin{pmatrix}\mathsf{I^t_0}\\ \mathsf{I^t_1}\\ \vdots\\ \mathsf{I^t}_{2k}\end{pmatrix} = \mathsf{A}$$

$$\frac{\partial}{\partial\alpha}\sum_{i=0}^{k}\sum_{j=0}^{k}\alpha_i\alpha_j\mathsf{E}X^{i+j} = (2\sum_{j=0}^{k}\alpha_j\mathsf{E}X^{i+j})_{i=0}^{k} = \begin{pmatrix}\alpha_0 & \alpha_1 & \dots & \alpha_k & 0 & \dots & 0\\ 0 & \alpha_0 & \alpha_1 & \dots & \alpha_k & \dots & 0\\ \vdots & & \vdots & & \vdots & \ddots & \vdots\\ 0 & \dots & 0 & \alpha_0 & \alpha_1 & \dots & \alpha_k\end{pmatrix}\begin{pmatrix}\mathsf{E}X^{0}\\ \mathsf{E}X^{1}\\ \vdots\\ \mathsf{E}X^{2k}\end{pmatrix} = \mathsf{E}$$

$$\frac{\partial}{\partial\alpha}\mathsf{F_i} = \mathsf{AC_1} - \mathsf{EC_1}^2\sum_{i=0}^{k}\sum_{j=0}^{k}\alpha_i\alpha_j\mathsf{I^t}_{i+j} =$$

$$= \mathsf{AC_1} - \mathsf{EC_1}^2\begin{pmatrix}\alpha_0 & \alpha_1 & \dots & \alpha_k\end{pmatrix}\begin{pmatrix}\alpha_0 & \alpha_1 & \dots & \alpha_k & 0 & \dots & 0\\ 0 & \alpha_0 & \alpha_1 & \dots & \alpha_k & \dots & 0\\ \vdots & & \vdots & & \vdots & \ddots & \vdots\\ 0 & \dots & 0 & \alpha_0 & \alpha_1 & \dots & \alpha_k\end{pmatrix}\begin{pmatrix}\mathsf{I^t_0}\\ \mathsf{I^t_1}\\ \vdots\\ \mathsf{I^t}_{2k}\end{pmatrix}$$

13. Для случая $\langle w\,, x_{\mathsf{i}}\rangle \geq 0$

$$\frac{\partial}{\partial w}\mathsf{F_i} = \frac{\partial}{\partial w}\mathsf{C_1}\sum_{i=0}^{k}\sum_{j=0}^{k}\alpha_i\alpha_j\int\limits_{-\infty}^{\langle w\,,x_{\mathsf{i}}\rangle}\frac{1}{\beta}t^{i+j}\frac{e^{-\frac{t-\alpha}{\beta}}}{(1+e^{-\frac{t-\alpha}{\beta}})^2}dt =$$

$$= \mathsf{C_1}\sum_{i=0}^{k}\sum_{j=0}^{k}\alpha_i\alpha_j\frac{\partial}{\partial w}\int\limits_{-\infty}^{\langle w\,,x_{\mathsf{i}}\rangle}\frac{1}{\beta}t^{i+j}\frac{e^{-\frac{t-\alpha}{\beta}}}{(1+e^{-\frac{t-\alpha}{\beta}})^2}dt =$$

$$= \mathsf{C_1} \sum_{i=0}^{k} \sum_{j=0}^{k} \alpha_i \alpha_j \tfrac{1}{\beta} \langle w\,, x_\mathsf{i} \rangle^{i+j} \frac{e^{-\frac{\langle w\,, x_\mathsf{i}\rangle - \alpha}{\beta}}}{\left(1+e^{-\frac{\langle w\,, x_\mathsf{i}\rangle - \alpha}{\beta}}\right)^2} \tfrac{\partial}{\partial w} \langle w\,, x_\mathsf{i} \rangle =$$

$$= \frac{e^{-\frac{\langle w\,, x_\mathsf{i}\rangle - \alpha}{\beta}}}{\beta\left(1+e^{-\frac{\langle w\,, x_\mathsf{i}\rangle - \alpha}{\beta}}\right)^2} x_\mathsf{i} \mathsf{C_1} \sum_{i=0}^{k} \sum_{j=0}^{k} \alpha_i \alpha_j \langle w\,, x_\mathsf{i} \rangle^{i+j}$$

$$\frac{\partial}{\partial w} \mathsf{F_i} = \frac{e^{-\frac{\langle w\,, x_\mathsf{i}\rangle - \alpha}{\beta}}}{\beta\left(1+e^{-\frac{\langle w\,, x_\mathsf{i}\rangle - \alpha}{\beta}}\right)^2} x_\mathsf{i} \mathsf{C_1} \begin{pmatrix} \alpha_0 & \alpha_1 & \dots & \alpha_k \end{pmatrix} \begin{pmatrix} \alpha_0 & \alpha_1 & \dots & \alpha_k & 0 & \dots & 0 \\ 0 & \alpha_0 & \alpha_1 & \dots & \alpha_k & \dots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \alpha_0 & \alpha_1 & \dots & \alpha_k \end{pmatrix} \begin{pmatrix} \langle w\,, x_\mathsf{i} \rangle^0 \\ \langle w\,, x_\mathsf{i} \rangle^1 \\ \vdots \\ \langle w\,, x_\mathsf{i} \rangle^{2k} \end{pmatrix}$$

14. Для случая $\langle w\,, x_\mathsf{i} \rangle < 0$

$$\tfrac{\partial}{\partial w} \mathsf{F_i} = \tfrac{\partial}{\partial w} \mathsf{C_1} \sum_{i=0}^{k} \sum_{j=0}^{k} \alpha_i \alpha_j \int\limits_{-\infty}^{\langle w\,, x_\mathsf{i}\rangle} \tfrac{1}{\beta} t^{i+j} \frac{e^{-\frac{t-\alpha}{\beta}}}{(1+e^{-\frac{t-\alpha}{\beta}})^2} dt =$$

$$= \mathsf{C_1} \sum_{i=0}^{k} \sum_{j=0}^{k} \alpha_i \alpha_j \tfrac{\partial}{\partial w} \int\limits_{-\infty}^{\langle w\,, x_\mathsf{i}\rangle} \tfrac{1}{\beta} t^{i+j} \frac{e^{-\frac{t-\alpha}{\beta}}}{(1+e^{-\frac{t-\alpha}{\beta}})^2} dt$$

Если $i + j$ четно, то функция симметрична относительно оси ординат.

$$\int\limits_{-\infty}^{\langle w\,, x_\mathsf{i}\rangle} \tfrac{1}{\beta} t^{i+j} \frac{e^{-\frac{t-\alpha}{\beta}}}{(1+e^{-\frac{t-\alpha}{\beta}})^2} dt = \int\limits_{-\infty}^{0} \tfrac{1}{\beta} t^{i+j} \frac{e^{-\frac{t-\alpha}{\beta}}}{(1+e^{-\frac{t-\alpha}{\beta}})^2} d - \int\limits_{0}^{\langle w\,, x_\mathsf{i}\rangle} \tfrac{1}{\beta} t^{i+j} \frac{e^{-\frac{t-\alpha}{\beta}}}{(1+e^{-\frac{t-\alpha}{\beta}})^2} dt$$

$$\tfrac{\partial}{\partial w} \int\limits_{-\infty}^{\langle w\,, x_\mathsf{i}\rangle} \tfrac{1}{\beta} t^{i+j} \frac{e^{-\frac{t-\alpha}{\beta}}}{(1+e^{-\frac{t-\alpha}{\beta}})^2} dt = -\tfrac{\partial}{\partial w} \int\limits_{0}^{\langle w\,, x_\mathsf{i}\rangle} \tfrac{1}{\beta} t^{i+j} \frac{e^{-\frac{t-\alpha}{\beta}}}{(1+e^{-\frac{t-\alpha}{\beta}})^2} dt = -\tfrac{1}{\beta} x_\mathsf{i} \langle w\,, x_\mathsf{i} \rangle^{i+j} \frac{e^{-\frac{\langle w\,, x_\mathsf{i}\rangle - \alpha}{\beta}}}{\left(1+e^{-\frac{\langle w\,, x_\mathsf{i}\rangle - \alpha}{\beta}}\right)^2}$$

Если $i + j$ нечетно, то это соответствует ситуации $\langle w\,, x_\mathsf{i} \rangle \geq 0$. Тогда

$$\tfrac{\partial}{\partial w} \mathsf{F_i} = \frac{e^{-\frac{\langle w\,, x_\mathsf{i}\rangle - \alpha}{\beta}}}{\beta\left(1+e^{-\frac{\langle w\,, x_\mathsf{i}\rangle - \alpha}{\beta}}\right)^2} x_\mathsf{i} \mathsf{C_1} \sum_{i=0}^{k} \sum_{j=0}^{k} \alpha_i \alpha_j (-1)^{[(i+j)\equiv 0 \pmod 2][\langle w\,, x_\mathsf{i}\rangle < 0]} \langle w\,, x_\mathsf{i} \rangle^{i+j}$$

15. В конечном итоге имеем

Обозначим $\mathsf{C}_{\mathsf{i+j}}^{\langle \mathsf{w}\,, \mathsf{x_i} \rangle} = (-1)^{[(i+j)\equiv 0 \pmod 2][\langle w\,, x_\mathsf{i}\rangle < 0]}$

$$\frac{\partial}{\partial w} \mathsf{F_i} = \frac{e^{-\frac{\langle w\,, x_\mathsf{i}\rangle - \alpha}{\beta}}}{\beta\left(1+e^{-\frac{\langle w\,, x_\mathsf{i}\rangle - \alpha}{\beta}}\right)^2} x_\mathsf{i} \mathsf{C_1} \begin{pmatrix} \alpha_0 & \alpha_1 & \dots & \alpha_k \end{pmatrix} \begin{pmatrix} \alpha_0 & \alpha_1 & \dots & \alpha_k & 0 & \dots & 0 \\ 0 & \alpha_0 & \alpha_1 & \dots & \alpha_k & \dots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \alpha_0 & \alpha_1 & \dots & \alpha_k \end{pmatrix} \mathsf{C}^{\langle \mathsf{w}\,, \mathsf{x_i} \rangle} \circ \begin{pmatrix} \langle w\,, x_\mathsf{i} \rangle^0 \\ \langle w\,, x_\mathsf{i} \rangle^1 \\ \vdots \\ \langle w\,, x_\mathsf{i} \rangle^{2k} \end{pmatrix}$$