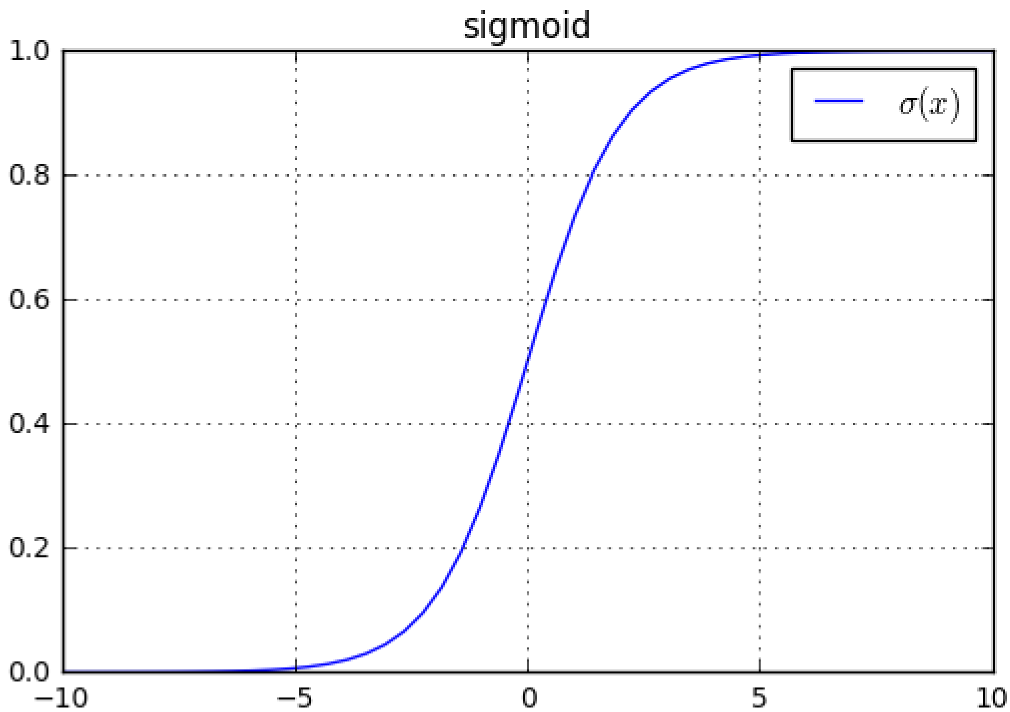
常用函数

一、 sigmoid

1. `sigmoid` 函数：

$$\sigma(x) = \frac{1}{1+\exp(-x)}$$

   - 该函数可以用于生成二项分布的 ϕ 参数
   - 当 x 很大或者很小时，该函数处于饱和状态。
     - 此时函数的曲线非常平坦，并且自变量的一个较大的变化只能带来函数值的一个微小的变化（即对自变量的变化不敏感）。



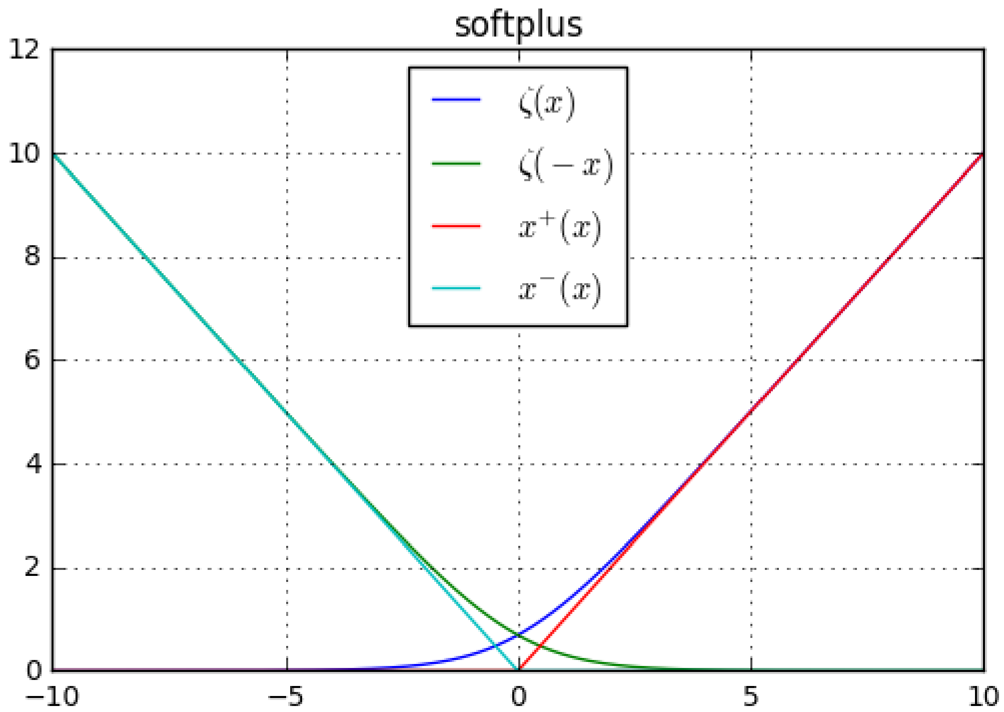
二、 softplus

1. `softplus` 函数：

$$\zeta(x) = \log(1+\exp(x))$$

   - 该函数可以生成正态分布的 σ^2 参数
   - 它之所以称作 `softplus`，因为它是下面函数的一个光滑逼近：

$$x^+ = \max(0,x)$$



2. `sigmoid` 和 `softplus` 函数的性质：

$$
\begin{gathered}
\sigma(x)=\frac{\exp(x)}{\exp(x)+\exp(0)}\\
\frac{d}{dx}\sigma(x)=\sigma(x)(1-\sigma(x))\\
1-\sigma(x)=\sigma(-x)\\
\log\sigma(x)=-\zeta(-x)\\
\frac{d}{dx}\zeta(x)=\sigma(x)\\
\forall x\in(0,1),\sigma^{-1}(x)=\log(\frac{x}{1-x})\\
\forall x>0,\zeta^{-1}(x)=\log(\exp(x)-1)\\
\zeta(x)=\int_{-\infty}^{x}\sigma(y)dy\\
\zeta(x)-\zeta(-x)=x
\end{gathered}
$$

其中 $f^{-1}(\cdot)$ 为反函数。

- $\sigma^{-1}(x)$ 也称作 `logit` 函数

3. 如果定义两个函数

$$
\begin{gathered}
x^+=\max(0,x)\\
x^-=\max(0,-x)
\end{gathered}
$$

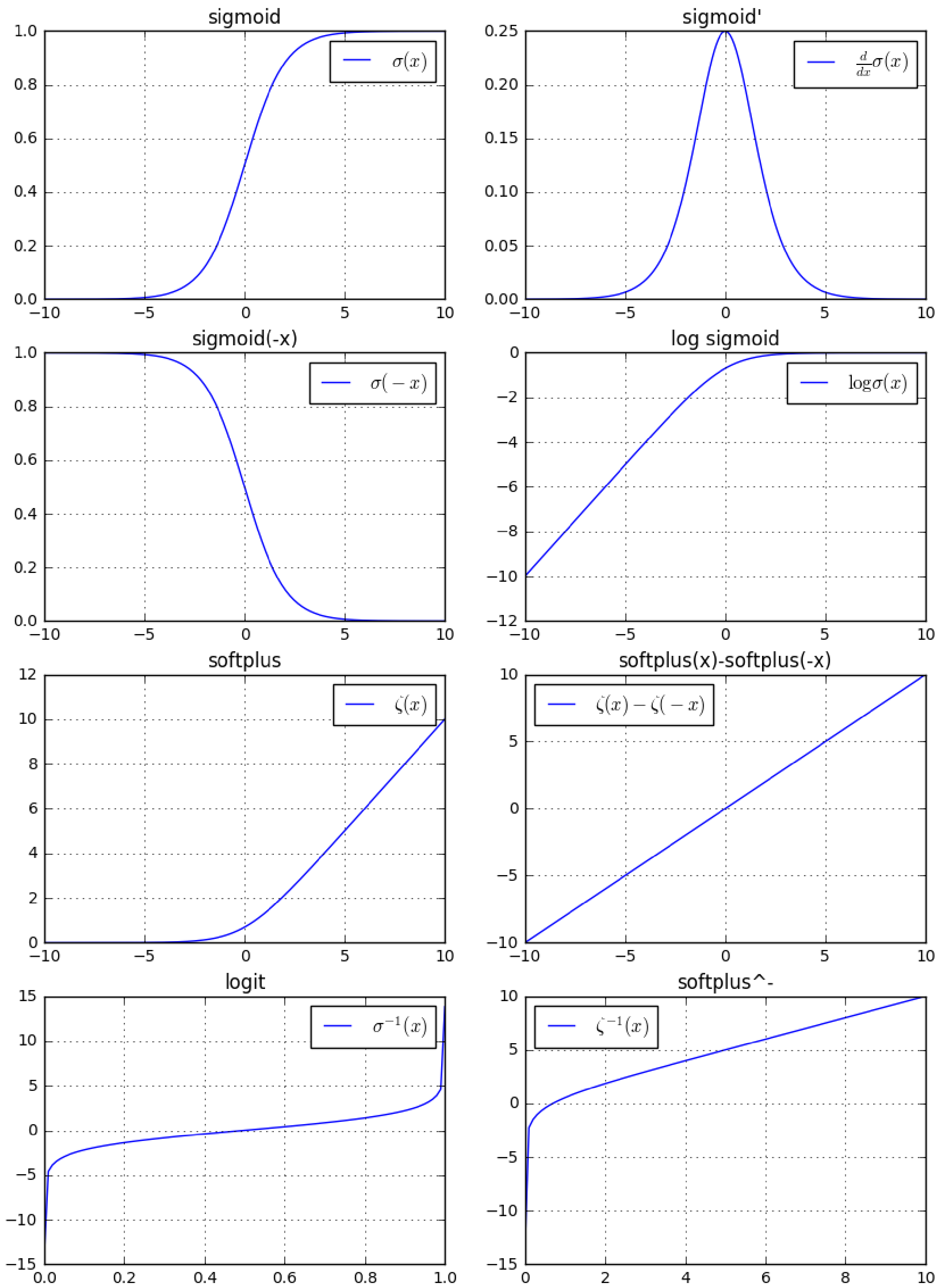
则它们分布获取了 $y=x$ 的正部分和负部分。

根据定义有：

$$x=x^+-x^-$$

而 $\zeta(x)$ 逼近的是 x^+， $\zeta(-x)$ 逼近的是 x^-， 于是有：

$$\zeta(x) - \zeta(-x) = x$$



三、Gamma 函数和贝塔函数

3.1 伽马函数

1. 伽马函数定义为：

$$\Gamma(x)=\int_0^{+\infty}t^{x-1}e^{-t}dt \quad ,x\in\mathbb{R}$$
$$or.\quad \Gamma(z)=\int_0^{+\infty}t^{z-1}e^{-t}dt \quad ,z\in\mathbb{Z}$$

   性质为：

   - $\Gamma(x+1)=x\Gamma(x)$，因此伽马函数是阶乘在实数域上的扩展。
     - 对于正整数 n 有：$\Gamma(n)=(n-1)!$
   - 与贝塔函数的关系：

$$B(m,n)=\frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)}$$

   - 对于 $x\in(0,1)$ 有：

$$\Gamma(1-x)\Gamma(x)=\frac{\pi}{\sin\pi x}$$

     则推导出重要公式：$\Gamma(\frac{1}{2})=\sqrt{\pi}$
   - 对于 $x>0$，伽马函数严格凹函数

2. 当 x 足够大时，可以用Stirling 公式来计算Gamma 函数值：

$$\Gamma(x)\sim\sqrt{2\pi}e^{-x}x^{x+1/2}$$

3.2 伽马分布

1. 伽马分布的概率密度函数为：

$$p(x;\alpha,\beta)=\frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x},\quad x>0$$

   记做 $\Gamma(\alpha,\beta)$

   - 期望 $\mathbb{E}[X]=\frac{\alpha}{\beta}$，方差$Var[X]=\frac{\alpha}{\beta^2}$

2. 性质：
   - 当 $\beta=n$ 时，为 `Erlang`分布
   - 当 $\alpha=1,\beta=\frac{1}{\lambda}$ 时，就是参数为 λ 的指数分布
   - 当 $\alpha=\frac{n}{2},\beta=\frac{1}{2}$ 时，就是常用的卡方分布
3. 伽马分布的可加性：设随机变量 $X_1,X_2,\cdots,X_n$ 相互独立并且都服从伽马分布 $X\sim\Gamma(x;)$

3.3贝塔函数

3.4 贝塔分布

1. 贝塔函数：

$$B(m,n)=2\int_0^{\frac{\pi}{2}}\sin^{2m-1}(x)\cos^{2n-1}(x)dx$$

- 当 $m,n>0$ 时，有：

$$B(m,n)=\int_0^1 x^{m-1}(1-x)^{n-1}dx$$

2. 当 m,n 足够大时，可以用Stirling 公式来计算贝塔函数值：

$$B(m,n)=\frac{\sqrt{(2\pi)}m^{m-1/2}n^{n-1/2}}{(m+n)^{m+n-1/2}}$$