

数据类型	FP64	FP32	TF32	FP16	BF16	FP8e5m2	FP8e4m3
符号位数	1	1	1	1	1	1	1
指数位数 (k)	11	8	8	5	8	5	4
尾数位数 (n)	52	23	10	10	7	2	3

- 最大值（最小值）

数据类型	FP64	FP32	TF32	FP16	BF16	FP8e5m2	FP8e4m3
最大值符号位 (s_2)	0	0	0	0	0	0	0
指数位 (e_2)	11111111110	11111110	11111110	11110	11111110	11110	1110
尾数位 (m_2)	111...1	111...1	111...1	111...1	1111111	11	111
$E = 2^e - (2^{k-1} - 1)$	1023	127	127	15	127	15	7
$M = \sum_{i=1}^n \frac{1}{2^i}$	$1 - \frac{1}{2^{52}}$	$1 - \frac{1}{2^{23}}$	$1 - \frac{1}{2^{10}}$	$1 - \frac{1}{2^{10}}$	$1 - \frac{1}{2^7}$	$1 - \frac{1}{2^2}$	$1 - \frac{1}{2^3}$
$max = (-1)^s 2^E (1 + M)$	1.798×10^{308}	3.403×10^{38}	3.401×10^{38}	65504.	3.390×10^{38}	57344.	240.
最小值符号位 (s_2)	1	1	1	1	1	1	1
$min = (-1)^s 2^E (1 + M)$	-1.798×10^{308}	-3.403×10^{38}	-3.401×10^{38}	-65504.	-3.390×10^{38}	-57344.	-240.

- 绝对最小值

数据类型	FP64	FP32	TF32	FP16	BF16	FP8e5m2	FP8e4m3
符号位 (s_2)	0	0	0	0	0	0	0
指数位 (e_2)	00000000000	00000000	00000000	00000	00000000	00000	0000
尾数位 (m_2)	000...01	000...01	000...01	000...01	0000001	01	001
$E = 1 - (2^{k-1} - 1)$	-1022	-126	-126	-14	-126	-14	-6
$M = \frac{1}{2^n}$	$\frac{1}{2^{52}}$	$\frac{1}{2^{23}}$	$\frac{1}{2^{10}}$	$\frac{1}{2^{10}}$	$\frac{1}{2^7}$	$\frac{1}{2^2}$	$\frac{1}{2^3}$
$value = (-1)^s 2^E M$	4.941×10^{-324}	1.401×10^{-45}	1.148×10^{-41}	5.960×10^{-8}	9.184×10^{-41}	1.526×10^{-5}	1.953×10^{-3}

- 其他值

数据类型	$+\infty$	$-\infty$	NaN
符号位 (s_2)	0	1	0 或 1
指数位 (e_2)	全 1	全 1	全 0
尾数位 (m_2)	全 0	全 0	非全 0