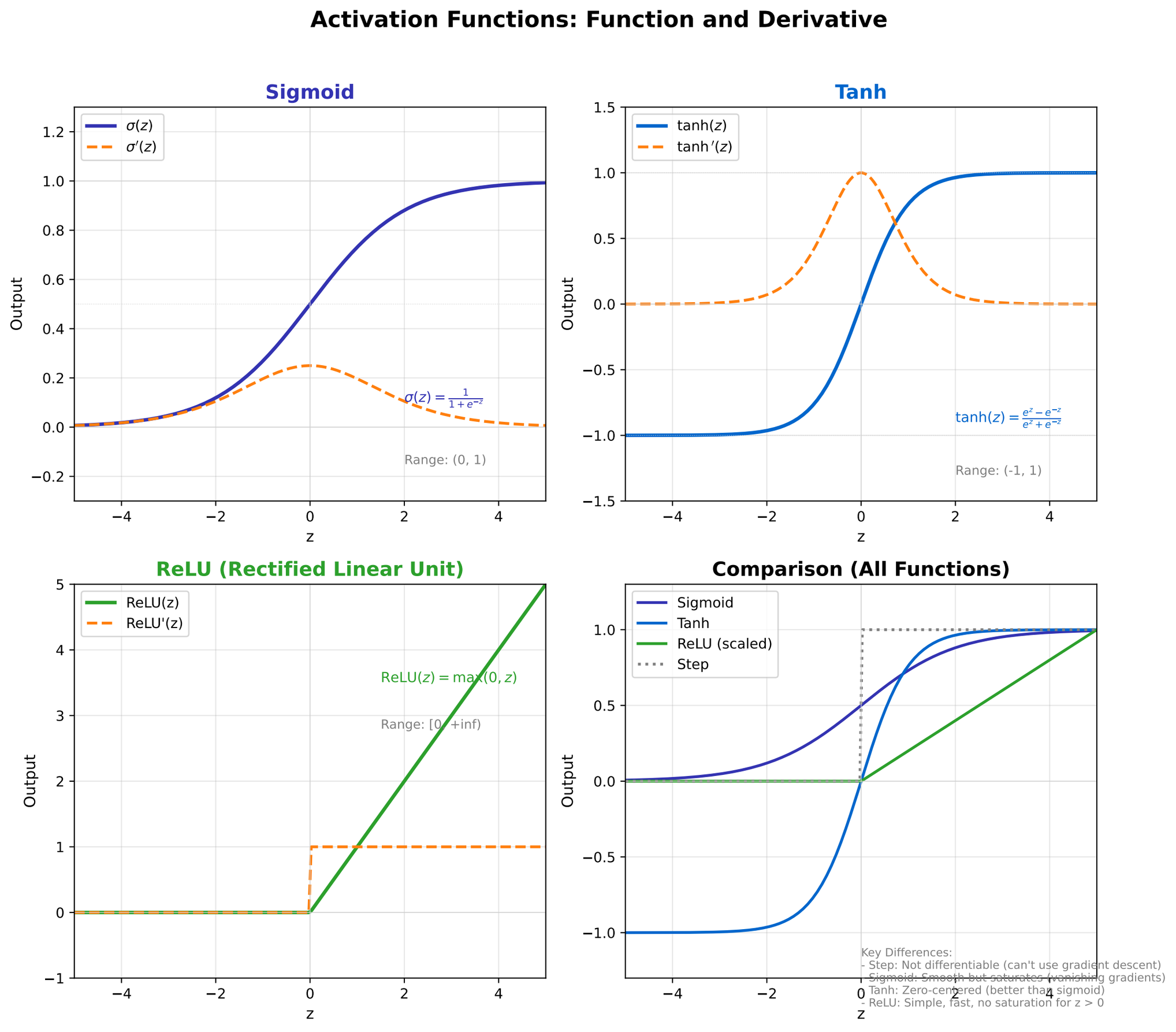
Activation Functions: Function and Derivative
Sigmoid
σ(z)
σ′(z)
$\sigma(z) = \frac{1}{1+e^{-z}}$
Range: (0, 1)
Output
z
Tanh
tanh(z)
tanh′(z)
$\tanh(z) = \frac{e^{z}-e^{-z}}{e^{z}+e^{-z}}$
Range: (-1, 1)
Output
z
ReLU (Rectified Linear Unit)
ReLU(z)
ReLU'(z)
ReLU(z) = max(0, z)
Range: [0, +inf)
Output
z
Comparison (All Functions)
Sigmoid
Tanh
ReLU (scaled)
Step
Key Differences:
- Step: Not differentiable (can't use gradient descent)
- Sigmoid: Smooth but saturates (vanishing gradients)
- Tanh: Zero-centered (better than sigmoid)
- ReLU: Simple, fast, no saturation for z > 0
Output
z