# LSTM 技术文档（Markdown 源码）

复制以下内容即可作为 .md 文件使用。

## 1 引言

LSTM 通过门控机制解决传统 RNN 的长程依赖问题，广泛应用于序列建模任务（机器翻译、语音识别、时间序列预测等）。

## 2 符号与维度

| 符号 | 含义 | 维度 |
| --- | --- | --- |
|  | 第 步输入 |  |
|  | 上一时刻隐藏状态 |  |
|  | 上一时刻细胞状态 |  |
|  | 输入权重矩阵 |  |
|  | 循环权重矩阵 |  |
|  | 偏置向量 |  |

其中 分别对应 **输入门、遗忘门、候选门、输出门**。

## 3 前向传播公式（单步）

给定 ，LSTM 按下列步骤计算：

1. **遗忘门**
2. **输入门 & 候选信息**
3. **细胞状态更新**
4. **输出门 & 隐藏状态**

## 4 反向传播（BPTT 核心）

记 ，，则：

* **细胞状态梯度**
* **门控梯度**（以遗忘门为例）
* **参数梯度**（以 为例）

其余门控与参数梯度同理可推。

## 5 PyTorch 代码范例

```python import torch.nn as nn

lstm = nn.LSTM(input\_size=d, hidden\_size=h, num\_layers=1, batch\_first=True)

# 输入: (batch, seq\_len, d)

out, (h\_n, c\_n) = lstm(x)