

时序数据实时处理方法

🕒 原	
👥 Stakeholders	
📌 Status	
📌 Type	
🕒 Created	@December 19, 2022 9:41 AM
🕒 Last Edited Time	@December 21, 2022 7:20 PM
👤 Last Edited By	

一、时序数据实时处理的发展路径

时序大数据（time series）= 批式大数据（batch）+ 流式数据（streaming）



二、时序数据实时分析计算：

关键技术一：面向复杂统计指标的增量计算

1. 数理统计算法的增量计算问题

简单算法 静态取数 容器类算法 复杂算法 CEP 等

2. 数理统计算法的可逆计算问题
3. 数理统计算法的乱序计算问题

关键技术二：面向时序数据处理的动态时间窗口

1. 时间窗口需提供滚动、滑动的漂移能力
2. 长周期时间窗口的动态精度控制
3. 基于弹性时间窗口的实时ADHoc查询

关键技术三：基于流的事件序列识别

1. 复杂事件处理CEP的增量匹配及数理统计问题

事件模式的增量匹配问题、叠加通用算法的增量统计问题

三、point by point：滤波器的信号处理是实时的，每读取一点数据分析一点数据，在分析一点数据的同时读取下一点的数据，然后重复上面的过程

四、系统对比

线性系统：同时满足叠加性与均匀性（又称为齐次性）的系统。所谓叠加性是指当几个输入信号共同作用于系统时，总的输出等于每个输入单独作用时产生的输出之和；均匀性是指当输入信号增大若干倍时，输出也相应增大同样的倍数。对于线性连续控制系统，可以用线性的微分方程来表示。不满足叠加性和均匀性的系统即为非线性系统。

因果系统：又称非超前系统 (nonanticipative system)即输出不可能在输入到达之前出现的系统。也就是说，系统 n 时刻的输出，只取决于系统 n 时刻以及 n 时刻之前的输入，而与 n 时刻之后的输入无关。系统的这种性质称为因果特性。

非因果系统(noncausal system)：指当前时刻的输出不仅取决于当前的输入，还取决于将来的输入的系统。

反因果系统(anticausal system)：当前时刻的输出仅取决于将来的输入的系统。

五、判定方法：

对于连续时间系统

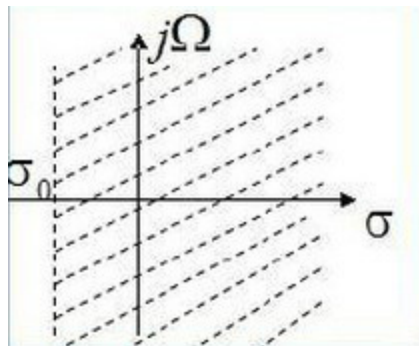
$t=t_1$ 的输出 $y(t_1)$ 只取决于 $t \leq t_1$ 的输入 $x(t \leq t_1)$ 时，则此系统为因果系统。特殊的，当该系统为线性移不变系统时：

(1) 时域判决：

系统的冲激响应函数 $h(t)$ ，在 $t < 0$ 的条件下， $h(t)=0$ ，则此系统为因果系统；如果系统的单位冲激响应在 $t > 0$ 时， $h(t)=0$ ，就说该系统是反因果的。

(2) S域判决：

系统函数的收敛域应该是s平面上某一收敛轴的右半平面。换句话说，系统函数的极点只能分布在s平面上收敛轴的左半平面



对于离散时间系统

(1) 时域判决：

$k=k_1$ 的输出 $y(k_1)$ 只取决于 $k \leq k_1$ 的输入 $x(k \leq k_1)$ 时，则此系统为因果系统。特殊的，当该系统为线性时不变系统时，系统的冲激响应函数 $h(k)$ ，在 $k < 0$ 的条件下， $h(k)=0$ ，则此系统为因果系统。即因果系统是激励加入之前不会出现响应的系统。

(2) 频域判决：

在Z域中，因果系统的判定为：

(1) 在 $H(z)$ 中不会出现 z 的正幂；

(2) $H(z)$ 的收敛域必在某圆外；

(3) 在下式中，只有 $m \leq n$ ：

$$H(z) = \frac{b_m z^m + b_{m-1} z^{m-1} + \dots + b_1 z + b_0}{a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0}$$