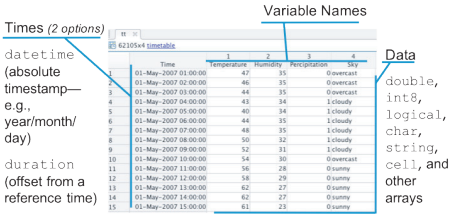
Предварительная обработка временных рядов в среде MATLAB

Этот обзор показывает общие случаи использования, но ни в коем случае не является исчерпывающим.
Значок » содержит ссылки на соответствующие разделы документации MATLAB, для более глубокого изучения.

Timetable

Тип данных timetable предназначен для организации и работы с данными временных рядов.

Компоненты timetable



Создание timetable »

```
tt = timetable(times, var1, var2,
      ... ,varN);
```

(Все переменные должны иметь одинаковое количество строк)

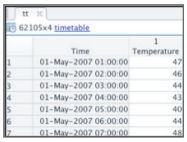
```
tt = table2timetable(t);
```

(Первая переменная t типа datetime или duration становится строками времени)

Обработка Timetable

Доступ к данным. На выходе получим один и тот же массив:

```
tt.Temperature
tt{:,'Temperature'}
tt{:,1}
```



Добавление новой переменной

```
tt.newVar = zeros(height(tt),1);
```

Изменение имени переменной

```
tt.properties.VariableNames =
newNames;
```

(Имена должны быть действительными идентификаторами MATLAB)

Совет: используйте matlab.lang.makevalidname для создания допустимых имен из потенциально недопустимых имен

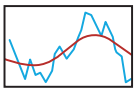
Изменение шага дискретизации с использованием retime

```
tt = retime(tt,newtimes,method);
```

метод используется для заполнения пробелов после retiming и имеет те же параметры, что и synchronize (**см. раздел "Merge Timetables"**).

Первичная обработка данных

Сглаживание данных »

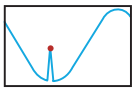


```
B = smoothdata(A,method);
```

Сглаживание зашумленных данных с помощью методов:

```
'movmean','movmedian','gaussian',
'lowess','loess','rlowess',
'rloess','sgolay'
```

Определение выбросов »

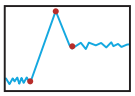


```
TF = isoutlier(A,method);
```

Идентификация выбросов с помощью методов:

```
'median','mean','quartiles',
'grubbs','gesd'
```

Определение точек изменения »



```
TF = ischange(A,method);
```

Поиск резких изменений с помощью методов:

```
'mean','variance','linear'
```

Объединение Timetable

Синхронизация нескольких timetable с общим вектором времени.

```
tt = synchronize(tt1,tt2,...,ttN);
```

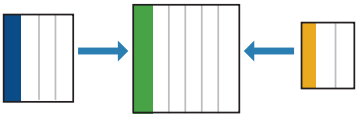
Синхронизация часто приводит к отсутствию точек данных (времен, в которые переменная не была измерена). `synchronize` поддерживает несколько методов корректировки данных для заполнения пропусков:

Заполнение: `'fillwithmissing','fillwithconstant'`

Интерполяция: `'linear','spline','pchip'`

Ближайшие соседи: `'previous', 'next','nearest'`

Агрегация: `'mean','min','max',@func,...`



Пропуски в данных

Нахождение пропусков

```
TF = ismissing(tt);
```

Заполнение пропусков

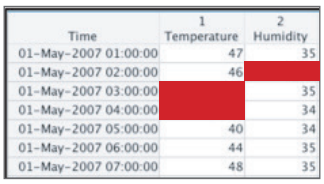
```
tt = fillmissing(tt,method);
```

Замените недостающие значения значениями, вычисленными из соседних точек с помощью методов:

```
'previous','next','nearest',
'linear','spline','pchip'
```

Удаление строк, содержащих пропуски

```
tt = rmmissing(tt);
```



Time	1 Temperature	2 Humidity
01-May-2007 01:00:00	47	35
01-May-2007 02:00:00	46	
01-May-2007 03:00:00		35
01-May-2007 04:00:00		34
01-May-2007 05:00:00	40	34
01-May-2007 06:00:00	44	35
01-May-2007 07:00:00	48	35

Большие данные

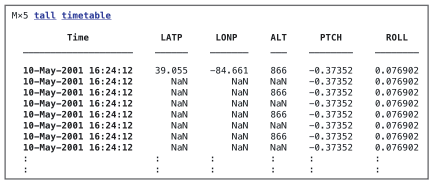
Tall arrays расширяют функции MATLAB для работы с данными, которые слишком большие для загрузки в память

Создание «tall» timetable:

```
% Создание datastore которое
% ссылается на данные
ds = datastore('*.csv');

% Создание tall table из
% datastore
t = tall(ds);

% Конвертация к timetable
tt = table2timetable(t);
```



M×5 tall timetable

Time	LATP	LONP	ALT	PTCH	ROLL
10-May-2001 16:24:12	39.055	-84.661	866	-0.37352	0.076902
10-May-2001 16:24:12	NaN	NaN	NaN	-0.37352	0.076902
10-May-2001 16:24:12	NaN	NaN	866	-0.37352	0.076902
10-May-2001 16:24:12	NaN	NaN	NaN	-0.37352	0.076902
10-May-2001 16:24:12	NaN	NaN	866	-0.37352	0.076902
10-May-2001 16:24:12	NaN	NaN	NaN	-0.37352	0.076902
10-May-2001 16:24:12	NaN	NaN	866	-0.37352	0.076902
10-May-2001 16:24:12	NaN	NaN	NaN	-0.37352	0.076902
⋮	⋮	⋮	⋮	⋮	⋮