

26. Matlab - Import a export XLS

Blaho Michal · MATLAB/Comsol

16.07.2010



V našom seriály pokračujeme importom a exportom súborov. Doteraz sme si povedali o importe a exporte [štandardných súborov Matlabu \(MAT-súborov\)](#) a [textových súborov](#). Dnes si povieme ako môžeme v Matlabe pracovať so súbormi tabuľkových procesorov hlavne tabuľkového procesora Excel.

Matlab podporuje prácu so súbormi niektorých tabuľkových procesorov. Funkcie spomenuté v dnešnom článku pracujú so súbormi tabuľkového procesora Microsoft Excel. Matlab obsahuje aj funkcie, ktoré pracujú so súbormi tabuľkového procesora Lotus 123. Tie však nájde užívateľ v príručke a sú veľmi podobné tým, ktoré si dnes spomenieme. V Matlabe 2008b, ktorý používame pri písaní tohto seriálu zatiaľ chýba podpora otvoreného formátu ods (tabuľkového procesora OpenOffice príp. NeoOffice). V unixových systémoch nemusia spomenuté funkcie fungovať správne.

Informácie o súbore

Matlab obsahuje príkaz **xlsinfo**, ktorá zistí či je súbor procesora Microsoft Excel. Príkaz vráti reťazec *Microsoft Excel Spreadsheet* ak je súbor čitateľný. V opačnom prípade vráti príkaz prázdny reťazec. Príkaz dokáže zistiť mená hárkov, ktoré súbor obsahuje.

```
>> [typ, desc] = xlsinfo('Zošit1.xls')
typ =
Microsoft Excel Spreadsheet
desc =
'Hárok1' 'Hárok2'
```

Pre porovnanie sme skúsili prečítať aj súbor procesora OpenOffice.

```
>> [typ, desc] = xlsinfo('Tabulka.ods')
typ =
''
desc =
Unreadable XLS file.
```

Export do XLS súboru

Pre export dát do môžete využiť funkciu **xlswrite**. Pre najjednoduchšie použitie funkcie musíme zadať meno súboru ako znakový reťazec a maticu s dátami. Ak sa súbor nenachádza v pracovnom adresári Matlab ho vytvorí. Matica by mala byť rozmeru MxN pričom platia isté obmedzenia. Hodnota M musí byť menšia ako 65536 a hodnota N menšia ako 256. Matica by mala byť číselného, alebo znakového dátového typu, prípadne typu cell (zmiešané dátové typy). Dáta budú zapísané do prvého hárku na pozíciu A1.

Meno hárku môžeme určiť ďalším parametrom v podobe znakového reťazca. Ak hárk neexistuje pridá sa na koniec hárkov. Matlab nás pritom upozorní textom *Warning: Added specified worksheet*. Dáta z matice môžeme uložiť do rôznych buniek v hárku. Bunky zadáme ako reťazec napríklad 'C2:H5', kde prvá bunka je ľavý horný a druhá bunka pravý dolný roh dát. Ak zasáme len jednu bunku bude to bunka ľavého horného rohu. Samozrejme ak je rozmer oblasti na zápis menší Matlab zapíše iba dáta, ktoré sa do danej oblasti zmestia. Ak je naopak oblasť väčšia Matlab do prázdnych buniek zapíše text #N/A. Ak sa zápis podarí Matlab vráti logickú 1 a v opačnom prípade logickú 0 spolu s chybovou správou.

```
>>M={'den','teplota';1,30;2,28;3,29;4,31}
>>xlswrite('Zosit.xls',M)
>>xlswrite('Zosit.xls',M,'Post1')
Warning: Added specified worksheet.
>>xlswrite('Zosit.xls',M,'Post1','B3:D8')
```

Výstup príkazov môžete vidieť na nasledujúcom obrázku. Znak #NaN je v slovenskej verzii označený ako #NEDOSTUPNÝ.

	A	B	C	D	E
1					
2					
3		den	teplota	#NEDOSTUPNÝ	
4		1	30	#NEDOSTUPNÝ	
5		2	28	#NEDOSTUPNÝ	
6		3	29	#NEDOSTUPNÝ	
7		4	31	#NEDOSTUPNÝ	
8		#NEDOSTUPNÝ	#NEDOSTUPNÝ	#NEDOSTUPNÝ	
9					
10					

Obr.1 Zápis do súboru

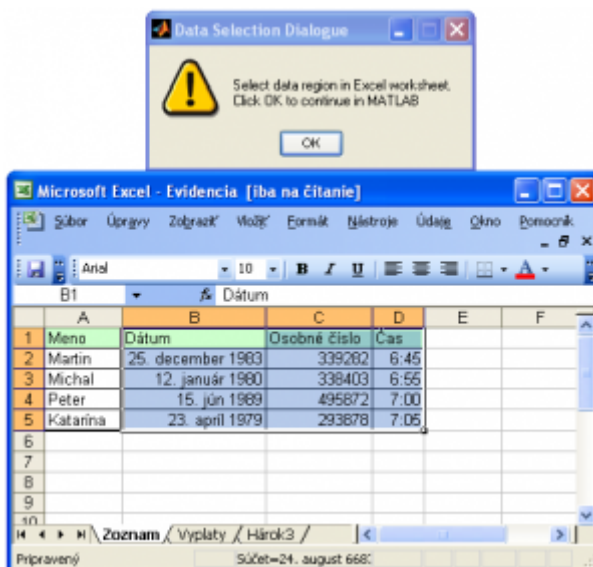
Import zo XLS súboru

Na čítanie dát z XLS súboru do Matlabu využívame príkaz **xlsread**. Funkcia je primárne určené na čítanie číselných dát. Preto po zadaní mena súboru vráti funkcia maticu čísel dátového typu double. Funkcia v základnom tvare ignoruje všetky nečíselné dáta mimo číselnej oblasti (napríklad popisy zľava alebo zhora). Ak sa v číselnej oblasti napriek tomu objavia nejaké nečíselné dáta sú nahradené číslom NaN.

Matlab umožňuje interaktívne vybrať dáta zo súboru. Stačí ako ďalší parameter zadať číselnú hodnotu -1. Matlab otvorí súbor, v ktorom označíte dátovú oblasť a svoju voľbu potvrdíte stlačením tlačidla OK. Na nasledujúcom obrázku je vidno interaktívny výber

dát. Dátum je v textovej podobe preto sa nebude nachádzať a numerickej časti výsledku príkazu.

```
>> num=xlsread('Evidencia.xls',-1)
num = 1.0e+005 *
3.392820000000000 0.399172813888889
3.384030000000000 0.399172887268519
4.958720000000000 0.399172918750000
2.938780000000000 0.399172951620370
```



Obr.2 Interaktívny výber

Podobne ako pri príkaze na export môžeme vybrať hárok a oblasť ako znakové reťazce pričom platia tie isté pravidlá, ktoré sme si spomenuli vyššie. Mená hárkov si zistíte pomocou spomínaného príkazu xlsinfo. Pre prečítanie nečíselných hodnôt, musíme pridať pri volaní funkcie jednu výstupnú premennú. Dáta budú v druhej premennej uložené v dátovom type cell.

```
>> [num,txt]=xlsread('Evidencia.xls','Zoznam')
num = 1.0e+005 *
3.392820000000000 0.399172813888889
3.384030000000000 0.399172887268519
4.958720000000000 0.399172918750000
2.938780000000000 0.399172951620370
```

```
txt =
'Meno' 'Dátum' 'Osobné číslo' 'Čas'
'Martin' '25. 12. 1983' ' '
'Michal' '12. 1. 1980' ' '
'Peter' '15. 6. 1989' ' '
'Katarína' '23. 4. 1979' ' '
```

Práca s dátumom z Excelu

Matlab načítava dátum z Excelu vo formáte v akom je uložený (numericky alebo textovo). Matlab aj Excel určujú dátum ako počet dní vzhľadom na nejaký referenčný

dátum. Matlab používa 0. január 0000 a Excel 1. január 1900. Preto treba dátumy po importovaní prepočítať, aby nám sedeli aj v Matlabe. S využitím funkcií **datenum** a **datestr** môžeme prepočítať hodnoty dátumov nasledovne:

```
>>excelD = xlsread('subor')
>>matlabD = datenum('30-Dec-1899') + excelD
>>datestr(matlabD,2)
```

Funkcia **datenum** konvertuje dátum na počet dní od referenčnej hodnoty Matlabu. Vstup do funkcie môže byť dátum ako reťazec znakov ('24-Oct-2003 12:45:07') prípadne vektor v tvare [Y M D H MS S] (rok, mesiac, deň, hodiny, minúty, sekundy). Funkcia **datestr** konvertuje dátum na jeho textovú podobu. Príkazu môžeme špecifikovať veľké množstvo želaných formátov, ktoré nájdete v tabuľke v manuály ku týmto príkazom. Najčastejšie však využijete reťazec ktorý bude obsahovať rok, mesiac, deň, hodiny, minúty, sekundy - 'yyyy.mm.dd HH:MM:SS' prípadne časť z neho s rôznymi oddeľovačmi medzi číslami (napr. čiarka, bodka, lomítka a pod.).

```
>>n = datenum(2001,12,19,18,0,0)
n = 7.312047500000000e+005
>> datestr(n, 'yyyy-mm-dd')
ans = 2001-12-19
>> datestr(n, 'yyyy/mm/dd HH:MM')
ans = 2001/12/19 18:00
```

Dnes sme si povedali o spracovaní XLS súborov pomocou príkazového riadka. Tieto príkazy sa vám môžu zísť pri hromadnom spracovaní dát pomocou Matlabu. Na import a export z XLS súborov v Matlabe môžu slúžiť aj ďalšie nastavby, ktoré už v tejto časti nespomíname. Ďalšie možnosti funkcií nájdete v užívateľskej príručke.

Spoluautorom článku je Gabriela Beňova.
