

情報リテラシ 第二

2Q 5c/6c(IL2) 木曜日

担当：地引

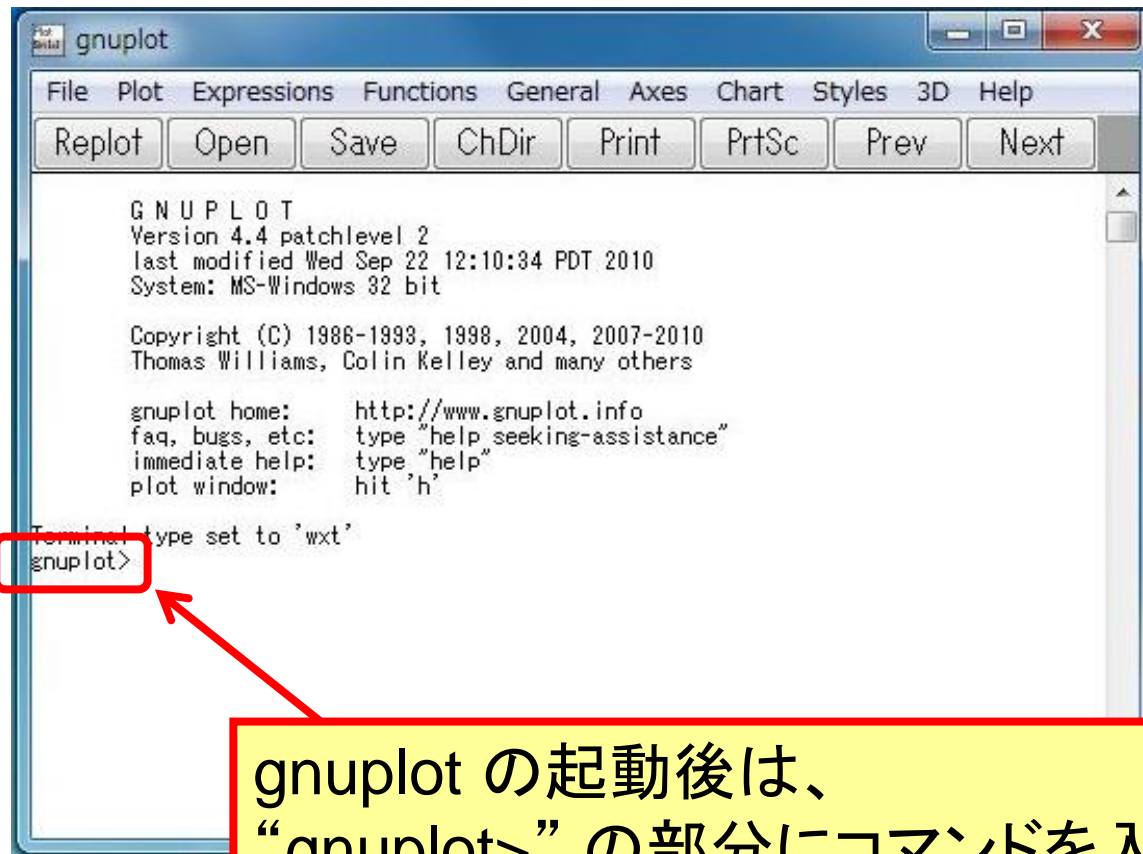
データの処理・加工 (特に可視化)

gnuplot とは

- 指定されたデータをもとにグラフを描画
- データの指定
 - 数式／数列 → どちらでも可
- Excel との違い
 - g: **3次元グラフ**を容易に描画できる。
 - g: 作成したグラフを容易に整形できる。
 - g: **プログラムとの連携**が容易にできるため、シミュレーション結果の表示などに使い易い。
 - E: 代表的な数値解析アルゴリズムが用意されている。
 - 回帰分析/共分散/線形計画など

gnuplot の起動

まずはターミナルを起動し、
“gnuplot” コマンドを入力して gnuplot を起動します。



```
gnuplot
File Plot Expressions Functions General Axes Chart Styles 3D Help
Replot Open Save ChDir Print PrtSc Prev Next
GNU PLOT
Version 4.4 patchlevel 2
last modified Wed Sep 22 12:10:34 PDT 2010
System: MS-Windows 32 bit

Copyright (C) 1986-1993, 1998, 2004, 2007-2010
Thomas Williams, Colin Kelley and many others

gnuplot home:      http://www.gnuplot.info
faq, bugs, etc:   type "help Seeking-assistance"
immediate help:   type "help"
plot window:      hit 'h'

Terminal type set to 'wxt'
gnuplot>
```

gnuplot の起動後は、
“gnuplot>” の部分にコマンドを入力します

gnuplot のカレント フォルダ

- Mac 上で起動したターミナルのカレント フォルダは、“/Users/ユーザ名” です（これは初期値です）。
- 同じく、Windows 上のコマンド プロンプトは、“C:¥Users¥ユーザ名” です（同上）。
- ターミナル上で gnuplot を起動した場合、gnuplot のカレント フォルダは、ターミナルと同じになります。
 - gnuplot はターミナルと同様、pwd, cd コマンドで、カレントフォルダを移動できますが、少し癖があります。
 - cd 'フォルダ名', フォルダ内を一覧できない他
 - gnuplot から、スクリプト ファイルやデータ ファイルの場所を正しく指定できるようになりましょう。

gnuplot の使い方 (1)

- 数式で指定されたグラフの描画
 - コマンドの入力: `gnuplot> plot 数式`
 - `plot`: 描画を指定するコマンド
 - 数式: 通常の数式をほぼ利用できる。
`sin(x)`, `2*x + 1` など

gnuplot の使い方 (2)

- ファイルに格納した数列をもとにした描画
 - 以下のフォーマットに従うファイルを作成する。

```
X1    y1
X2    y2
X3    y3 } 一つの座標は1行に書く
      ⋮
```

- `gnuplot > plot “データファイル”`
 - “と”を忘れない。
 - 例) `plot “作業フォルダ/gp_sample.data”`

gnuplot を起動したターミナルのレントフォルダが、gnuplot のレントフォルダになります。

要注意

フォルダの階層とその表現を
理解できていますか？

自分が見ているフォルダ(カレントフォルダ)と
対象ファイルが存在するフォルダの位置関係は？

gnuplot の使い方 (3)

- 空行の挿入

- データの塊毎にグラフを描画し、各グラフを結合

X11	y11	Z11
-----	-----	-----

X12	y12	Z12
-----	-----	-----

X13	y13	Z13
-----	-----	-----

空行で区切られた塊は
一つのグラフとして描画

X21	y21	Z21
-----	-----	-----

X22	y22	Z22
-----	-----	-----

X23	y23	Z23
-----	-----	-----

⋮

各塊の対応する点を
接続した平面として
描画

gnuplot の使い方 (4)

- ラインスタイル(線種)の変更
 - `gnuplot> plot “データファイル” with line_style`
 - 代表的な `line_style`: `lines`, `points`, `linespoints`, `dots`, `impulses`, `errorbars`, `boxes` ほか
- 複数の数式/データファイルの表示
 - `gnuplot> plot 数式 with line_style, 数式 with line_style`
 - 複数の数式指定をカンマ(,)で区切る。
- gnuplot には、これら以外にも様々な機能がある。
 - 教材に載せたリンクを参考にしよう。

gnuplot の使い方 (5)

3 次元の数式の入力

– $z = f(x, y)$ の場合

- `gnuplot> splot f(x, y)`

ex) `gnuplot> splot x*y`

– パラメータ表示の場合 ($x = f(u, v)$, $y = g(u, v)$, $z = h(u, v)$)

- `gnuplot> set parametric`

- `gnuplot> splot f(u, v), g(u, v), h(u, v)`

ex) `gnuplot> set parametric`

`gnuplot> splot cos(u)*cos(v), cos(u)*sin(v), sin(u)`

– パラメータ表示を止める場合

- `gnuplot> set noparametric`

演習 (1)

gnuplot を用いて、次のグラフを描画しなさい。

– $y = x \cdot \sin(x)$, $y = 3x^2$, $z = x \cdot y$

– パラメータ表示された

$\cos(u) \cdot \cos(v)$, $\cos(u) \cdot \sin(v)$, $\sin(u)$

を描画しなさい。

– 以下の手順により、

ファイル内に格納されたデータをグラフ化しなさい。

- ターミナル起動時のカレントフォルダに作業フォルダを作成 (名前は何でもよい/英数文字のみ)
- 上で作成した作業フォルダへ、教材の“gnuplot 用データファイル”からサンプルデータ (gp_sample.data) を保存 ⇒ Chrome が良い?
- 保存したサンプルデータ内のデータを描画

出力の保存

- 出力(グラフ)を再利用するために、ファイルへ保存する。
 - `gnuplot > set terminal` **出力形式 オプション**
 - 出力形式: pdf → 標準的な電子文書用データ形式
 - `gnuplot > set terminal pdf`
- 出力結果は、画面ではなくファイルへ格納される。
 - `gnuplot > set output` **“ファイル名.pdf”**
 - ファイルの保存先に注意 (**gnuplot のカレントフォルダは?**)
 - 例) `set output “作業フォルダ/○○○.pdf”`
 - MacOS の gnuplot では、gnuplot の終了時にファイルへの保存が実行されます。

フォルダの階層とその表現を 理解できていますか？

自分が見ているフォルダ(カレントフォルダ)と
対象ファイルが存在するフォルダの位置関係は？

gnuplot の利用例

凡例を付けない

```
gnuplot> set nokey
gnuplot> set terminal pdf
gnuplot> set output "work/graph.pdf"
gnuplot> plot [-30:30][-20:20] x*sin(x)
```

x 座標は -30 ~ 30,
y 座標は -20 ~ 20 までを表示

一連の gnuplot コマンドをファイルに保存しておく、
データの処理/可視化を行なう gnuplot プログラムとして扱える。

演習 (2)

- “情報リテラシ第二” → “テーマ1 データの処理と加工” → “実習に用いるデータセット” より、実習データをダウンロードし、作業フォルダに展開する。
 - 事前に、作業フォルダを自身のホームフォルダ以下に作成しておく。
- “情報リテラシ第二” → “テーマ1 データの処理と加工” → “実習A: 図5と図6の折れ線グラフ” を作成しなさい。
 - まずは、図5の折れ線グラフを作成します。
 - ターミナルを開き、展開した実習データの plot フォルダに移動します。
 - コマンドラインより、“gnuplot fig5-all-etime-size.plot” を実行すると、images フォルダ内に PDF 形式のグラフが作成されます。
 - 次に、図6のグラフを作成します。
 - fig5-all-etime-size.plot を参考に、エディタを用いて、図6のグラフを作成するための fig6-etime-size-M.plot を作成しましょう。

カレントフォルダを中心とした階層関係に気を配りましょう(迷子にならないように)。

未知の問題に取り組むとは？

(手掛かりがないとは、どういうこと？)

データの特徴を掴むには

- データ /cnm/ration-join.data は、そのまま表示しても特徴をよく掴めない。特徴をうまく掴めるような工夫が欲しい。
 - データの分布が、極端に偏っているように見える。
 - 相互関係では、指数法則の影響を受ける場合が多い
(強い存在が益々 / 指数的に強くなる: スケールフリー)。



- x, y 軸を対数表示に変えてみる。
 - `gnuplot > set logscale 軸`

グラフの出力形式について

- pdf 形式：図形を数式で表現
 - 一種のプログラムと言える。
 - 曲線は滑らかに表示される。
 - 拡大/縮小しても、画質は影響されない。
 - 描画するデータが増えると、処理が遅くなる(pdf ファイルのサイズも大きくなる)
- ビット マップ形式：図形を点の集まりとして表現
 - 点(画素)の細かさには限界があるため、曲線の描画や拡大/縮小時に画質が落ちる(掠れたように見える)。
- 出力形式をビット マップ形式に変える。
 - `gnuplot > set terminal png`

演習 (3)

- “情報リテラシ第二” → “テーマ1 データの処理と加工” → “実習B: 図3と図7の散布図” を作成しなさい。
 - まずは、図3の散布図を作成します。
 - 展開した実習データの plot フォルダに移動します。
 - “gnuplot fig3-cnm-ratio-joins.plot” により作成されるグラフの y 軸が対数目盛となるように fig3-... を修正しましょう。
 - 次に、図7のグラフを作成します。
 - 修正した fig3-cnm-ratio-joins.plot を参考に、エディタを用いて、図7のグラフを作成するための fig7-hn-ratio-joins.plot を作成しましょう。
 - エディタで fig3-... を開き、**修正する前に** fig7-... という名前で保存しておくと効率的です。
- .plot ファイル内に記述されている gnuplot のコマンドと、実際に描画されたグラフとの確認を試みよう。

カレントフォルダを中心とした階層関係に気を配ろう。