

第6回 数値データの可視化

- 目的: データを図で表現すると説得力が出てくる。表計算ソフト(EXCEL)でもデータからグラフを作ることができる。ここでは、より多機能(だが、やや面倒)なgnuplotというソフトを使ってみよう。



- 内容
 - データを作る
 - 基本的な作図
 - オプションを設定する
 - グラフを保存する
 - 宿題

データを作る

- テキストエディタにて下のデータを打ち込み temperature.dat という名前で保存する。
 - 半角で入力し、隣のデータとはTabで区切る。
 - Num Lockを押すとテンキーから入力できる。

1	1.5	5.2	-2
2	1.7	5.5	-1.8
3	4.5	8.8	0.5
4	10.1	14.8	5.7
5	14.9	19.5	10.8
6	18.3	22	15.3
7	22.1	25.7	19.3
8	24.1	27.9	21.2
9	20.4	24.1	17.2
10	14.8	19.1	10.8
11	9.1	13.4	4.9
12	4.3	8.3	0.6

データの区切りはTabあるいはスペース

データは、1971～2000の月別の平均気温、最高気温、最低気温である。

仙台管区気象台のウェブサイト

基本的な作図

- 端末エミュレータからgnuplot(グニユプロット、グヌプロット)を起動
c2ja0rat(111)% gnuplot

```
G N U P L O T
Unix version 3.7 .....
```

- コマンドを入力する。
plot *option, option,*

以下の例に従ってグラフを順次作成してみよう

```
gnuplot> plot "temperature.dat"
```

⇒temperature.dat ファイルをプロット(X軸のデータは1列目、Y軸のデータは2列目)

```
gnuplot> plot "" with lines
```

⇒""は、直前のファイルを指定。with linesは折れ線グラフ

```
gnuplot> plot "" with errorbars, "" with lines linetype 1 linewidth 2
```

⇒ with errorbarsは3,4列目のデータを縦線で結ぶ(誤差線)

⇒, は重ねて次のプロットをする。linetype は色、linewidth は線の太さ

```
gnuplot> plot "" notitle with errorbars, "" smooth cspline notitle linetype 1 linewidth 2
```

⇒ notitleはデータのタイトルを表示しない。

⇒ smooth cspline はデータを滑らかに結ぶ(補完)

細かく設定する

- 細かい設定は set を使う。以下をやってみよう。

```
gnuplot> set title "Monthly average temperature¥n SENDAI, 1971-2000"
```

⇒グラフのタイトルをつける。¥nは改行

```
gnuplot> set xlabel "Month" ⇒ X軸のラベルをつける。
```

```
gnuplot> set xrange [0.5:12.5] ⇒ X軸の範囲を0.5~12.5にする。
```

```
gnuplot> set xtics 1 ⇒ X軸のメモリの増分を設定する。
```

```
gnuplot> set ylabel "Temperature¥n(degree)" ⇒ Y軸のラベルをつける。
```

```
gnuplot> replot ⇒グラフを再描画する。
```

グラフを保存する

- 表示されたグラフを保存するには:

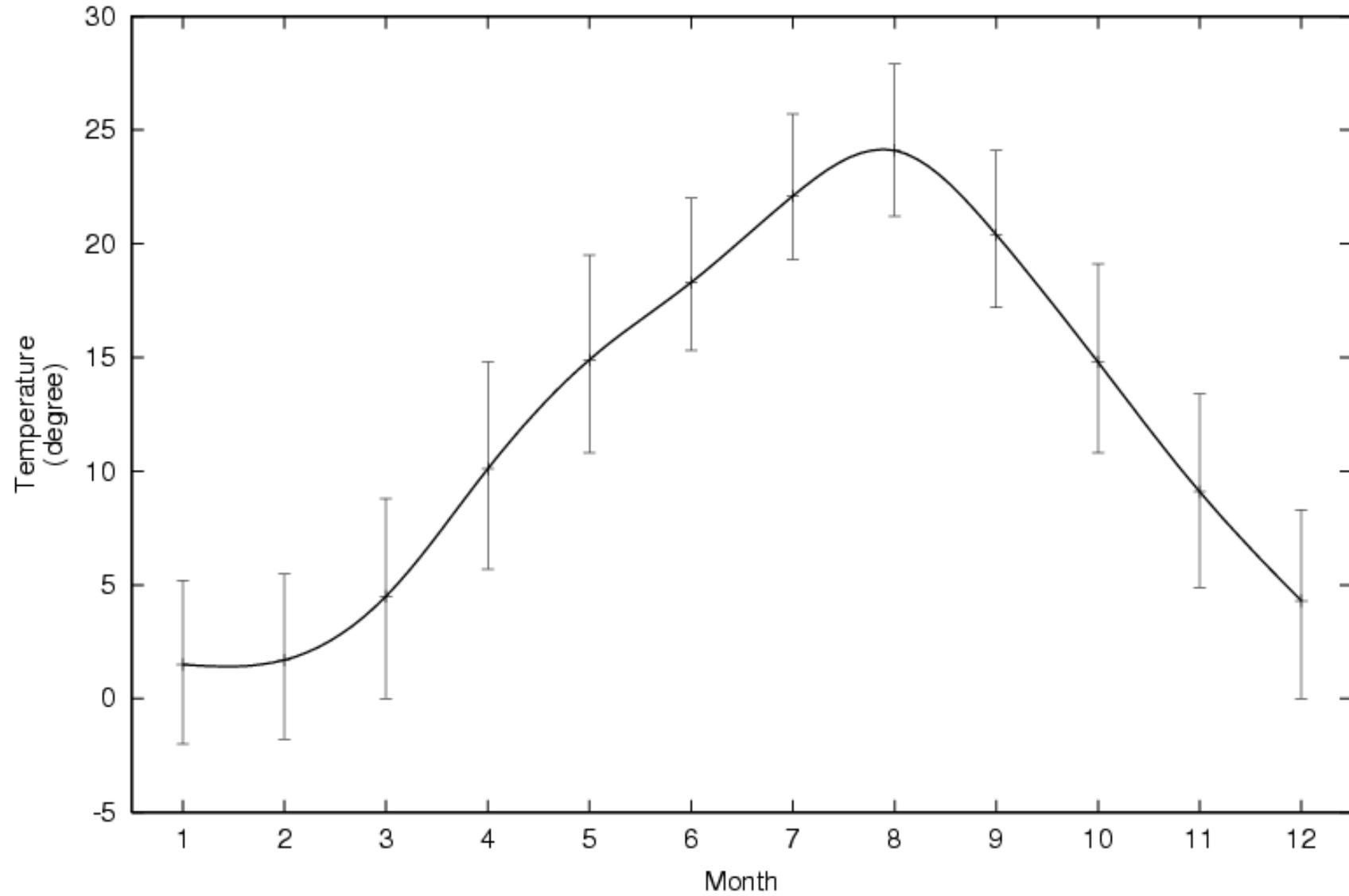
gnuplot> set terminal postscript	⇒ポストスクリプト形式(*)にする。
gnuplot> set output "kion.ps"	⇒保存するときのファイル名を指定。
gnuplot> replot	⇒描画したものを保存する。
gnuplot> set terminal x11	⇒標準の形式にする。
gnuplot> replot	⇒通常の方法で表示される(確認)
gnuplot> exit	⇒gnuplotの終了

保存されたグラフは コマンドプロンプトから
c2ja0rat(149) ghostview kion.ps
で見ることができる。

ポストスクリプト(PostScript)形式(*):

一種のプログラミング言語である。印刷あるいは表示する時に、文字や図形などの要素を点の集まりではなく、曲線で表す。文字を曲線の情報で持つことで、拡大してもギザギザができずに滑らかに表示される。

Monthly average temperature
SENDAI, 1971-2000



宿題

- gnuplotで二次元のグラフを作成し、次回の講義までにポストスクリプトファイルで保存しておく。
データの素材はwebから探し、最低でもプロットが10点になるようにする(ex. 年間データ10年分)。また、データの出典(webサイト名)と、学籍番号、名前をタイトルに必ず入れる。
gnuplotの使い方はgnuplot>help で出てくるが。英語である。日本語の解説のサイトを検索すると、数多くのwebサイトがあるので参考にすると良い。

例:

- ある県の人口の推移
- 野球またはサッカーのチームの順位の変動
- あるメーカーのあるものの売上の推移(例えば、ホンダのCIVICの売上台数)
- 日本の米の生産額の推移