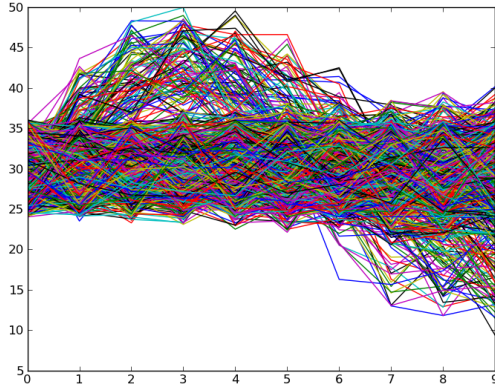


SVD ile Zaman Serisi Kumeleme

Tekil Deger Ayristirma (Singular Value Decomposition -SVD-) ile bir veri madenciligi ornegi gorecegiz. Ornek olarak [1] adresinde tarif edilen / paylasilan zaman serisini kullandik. Once veriyi grafikledik,



Verinin tamamı kullanılmadı, serinin ilk 10 noktasını aldık, ve grafiğe bakınca iki tane ana seri olduğunu görüyoruz.

```
import numpy as np
from pylab import *

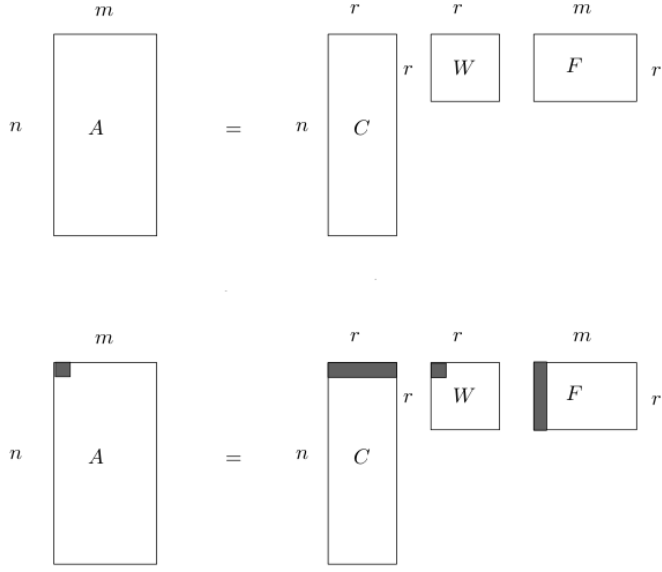
data = np.genfromtxt("synthetic_control.data", dtype=float)

print data.shape

for t in data[:,0:10]:
    plot(t); hold(True)

show()
```

Peki bu serileri nasıl otomatik olarak kümeleyerek bulurduk / birbirinden ayırtırdık? *Lineer Cebir Ders 29*'da SVD'nin matematigini isledik. SVD bir matris A üzerinde ayırtırma yapar, ve A herhangi boyutta, türde bir matris olabilir.



Ayrıştırmanın $A = CWF$ sonucunu verir, burada C , ana matris ile aynı miktarda satıra sahiptir, F aynı miktarda kolona sahiptir. Ayrıştırma sonrası A 'nin kertesesi (rank) ortaya çıkar, eğer tüm A kolonları birbirinden bağımsız ise, o zaman $r = m$ olacaktır, ama kolonların bazıları mesela aynı ölçümü değişik katlarda tekrarlıyor ise, o zaman matriste tekillik vardır, ve bu durumda $r < m$ olur, ve ortadaki W matrisi $r \times r$ olduğu için beklenenden daha ufak boyutlarda olabilir.

[1] http://kdd.ics.uci.edu/databases/synthetic_control/synthetic_control.data.html