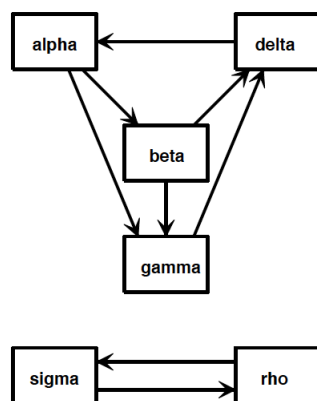


Problems 1 and 2 of this homework include some initial practice materials. Problem 3 is a “tell me about yourself” survey, which you can find in Canvas under “quizzes”. It is worth 10pts and is graded automatically by Canvas.

- (1) (10pts) A small model of an internet with 6 pages is



- Find the *connectivity* matrix G for this network and implement it as a sparse array in Matlab.
- Use a transition probability of $p = 0.85$ to compute the PageRank for this web. You can use the function `pagerank.m` from the NCM code files on the coursepage with the cell array $U = \{\text{http://www.alpha.com}; \text{http://www.beta.com}; \dots\}$.
- The algorithm is in part based on estimating the eigenvector corresponding to $\lambda = 1$ of the transition matrix A . Show that this eigenvalue exists, by finding the special vector y so that $y^T A = y^T$. Explain why this means that PageRank's eigenvector $Ax = x$ also exists.

- (2) (10pts) (adapted from Problem 5.8 in [1]). Below are 25 observations y_k , at equally spaced values of t

```
t = (1:25)';
y = [ 5.0291  6.5099  5.3666  4.1272  4.2948
      6.1261 12.5140 10.0502  9.1614  7.5677
      7.2920 10.0357 11.0708 13.4045 12.8415
      11.9666 11.0765 11.7774 14.5701 17.0440
      17.0398 15.9069 15.4850 15.5112 17.6572];
y = y';
```

- Fit the data with a straight line, $y(t) = \beta_1 + \beta_2 t$, and plot the residuals, $y(t_k) - y_k$. You should observe that one of the data points has a much larger residual than the others. This is probably an outlier.
 - Discard the outlier, and fit the data again by a straight line. Plot the residuals again. Do you see any pattern in the residuals?
 - Fit the data, with the outlier excluded, by a *linear* model of the form $y(t) = \beta_1 + \beta_2 t + \beta_3 \sin t$.
 - Evaluate the third fit on a finer grid over the interval $[0, 26]$. Plot the fitted curve, using line style '-', together with the data, using line style 'o'. Include the outlier, using a different marker, '*'.
 - Fit the data, with the outlier excluded, by a *nonlinear* model of the form $y(t) = \beta_1 + \beta_2 t + \beta_3 \cos(\beta_4 - t)$. Include the results in the plots and comment on the value of β_4 . For the fitting process you can choose an initial value of $\beta_4^{(0)} = 0$.
- (3) (10pts) Go to Canvas quizzes and take the survey “First Day Survey: Tell Me About Yourself #FinAid”. Note that you have unlimited attempts, except you must be done by Friday, April 14, at 11:59pm).