

1. Let S be the system

$$\begin{aligned} x - y &= 1 \\ -x + y &= 1 \end{aligned}$$

(a) What is the output of Gauss-Jordan elimination on system S ? Show the steps of the algorithm.

Solution: After eliminating x , the system becomes $0 = 2$. The algorithm outputs unsatisfiable.

(b) What does Gradient Descent converge to on initializations $(1, 1)$ and $(1, -1)$ at rate $\rho = 0.01$?

Solution: The state evolves as $\mathbf{x}' = \mathbf{x} - 2\rho A^T(A\mathbf{x} - b) = (I - 2\rho A^T A)\mathbf{x} + 2A^T b$. For this system,

$$\rho A^T A = \begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix} \quad \text{and} \quad 2A^T b = 0.$$

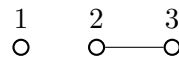
$(1, 1)$ is an eigenvector of $A^T A$ with eigenvalue 0. If gradient descent starts at this state it never moves. It converges to $(1, 1)$.

$(1, -1)$ is an eigenvector of $A^T A$ with eigenvalue 2. Each step of Gradient Descent shrinks it by a factor of $1 - 4\rho$. When $\rho = 0.01$, it converges to zero.

(c) What does Gradient Descent converge to on initialization $(2, -3)$ at the same rate? (Use part (b).)

Solution: $(2, -3) = -\frac{1}{2}(1, 1) + \frac{5}{2}(1, -1)$. The first component is stable under gradient descent. The second one eventually vanishes. The algorithm will converge to $(-1/2, -1/2)$. The coefficients $-1/2$ and $5/2$ were derived by projecting $(2, -3)$ onto the eigenvectors, e.g., $(2, -3) \cdot (1, 1) / \|(1, 1)\|^2 = -1/2$.

2. Let S be the adjacency matrix of the following graph:



(a) Write the first four iterates of power iteration on input S with generic initialization (a, b, c) :

Solution: The matrix S annuls the first entry and swaps the other two. The iterates are

$$\mathbf{x}_0 = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \quad \mathbf{x}_1 = \begin{bmatrix} 0 \\ c \\ b \end{bmatrix} \quad \mathbf{x}_2 = \begin{bmatrix} 0 \\ b \\ c \end{bmatrix} \quad \mathbf{x}_3 = \begin{bmatrix} 0 \\ c \\ b \end{bmatrix}.$$

These are the unnormalized iterates. The normalized ones are scaled by $1/\sqrt{b^2 + c^2}$.

(b) Use part (a) to deduce the eigenvalues of S from largest λ_1 to smallest λ_3 . Explain your reasoning.

Solution: When $a = 1$ and $b = c = 0$ power iteration converges to zero, so zero must be one of the eigenvalues. When b and c are different, say $b = 1$ and $c = 0$, then it keeps flipping between $(0, 1, 0)$ and $(0, 0, 1)$. As the length is preserved, S must have spectral norm 1. As power iteration does not stabilize, the eigenvalues must be -1 and 1 . Therefore $\lambda_1 = 1$, $\lambda_2 = 0$, and $\lambda_3 = -1$.

(c) What does subspace iteration on input $S' = S - \lambda_3 I$ initialized with the standard basis converge to?

Solution: The matrix S' is

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

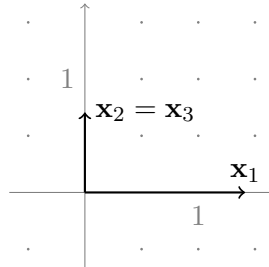
The vector $\mathbf{x}_1 = (1, 0, 0)$ doesn't change in the iterates. The vectors $\mathbf{x}_2 = (0, 1, 0)$ and $\mathbf{x}_3 = (0, 0, 1)$ both become $(1/\sqrt{2})(0, 1, 1)$ after one step and don't change in further iterates.

(d) Calculate the spectral decomposition of S' . Plot the vectors $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ that realize S' as dot products.

Solution: The eigenvalues of S' are 2, 1, 0. By part (c) its spectral decomposition is $A^T A$ with

$$A = \begin{bmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

The embedding of the graph are the columns of A .



3. The function f_n over domain $\{+, -\}^n$ counts the number of consecutive $+-$ patterns in the input. For example, $f_7(\underline{+}, \underline{-}, +, \underline{+}, \underline{-}, \underline{+}, \underline{-}) = 3$ and $f_5(\underline{-}, \underline{-}, \underline{-}, \underline{-}, \underline{+}) = 0$.

(a) What is the function $f_{n+1}(x_1, \dots, x_{n+1}) - f_n(x_1, \dots, x_n)$?

Solution: The two functions count the number of consecutive $+-$ patterns in the first $n + 1$ and n positions, respectively. Their difference indicates the presence of $+-$ at positions n and $n + 1$. This is

$$\text{point}_{+-}(x_n, x_{n+1}) = \frac{1 + x_n}{2} \cdot \frac{1 - x_{n+1}}{2} = \frac{1}{4} + \frac{1}{4}x_n - \frac{1}{4}x_{n+1} - \frac{1}{4}x_n x_{n+1}.$$

(b) Calculate $\hat{f}_8(\emptyset)$, $\hat{f}_8(\{2\})$, and $\hat{f}_8(\{5, 6\})$. (Use part (a).)

Solution: Expanding the recursion in part (a),

$$f_8 = \text{point}_{+-}(x_1, x_2) + \text{point}_{+-}(x_2, x_3) + \dots + \text{point}_{+-}(x_7, x_8).$$

The constant term $\hat{f}_8(\emptyset)$ is the sum of the constants of all seven point functions, namely $7 \cdot 1/4 = 7/4$.

The term $\hat{f}_8(\{2\})$ is the coefficient in front of x_2 . It is $1/4$ in the first point function and $-1/4$ in the second one. They cancel out and $\hat{f}_8(\{2\})$ is zero.

The monomial $x_5 x_6$ only appears in $\text{point}_{+-}(x_5, x_6)$ with coefficient $-1/4$, so $\hat{f}_8(\{5, 6\}) = -1/4$.