DSE 220: Machine learning

Worksheet 12 — Solutions

1. *Singular values versus eigenvalues.*

   (a) We have expressed M using its singular value decomposition: $M = U\Lambda V^T$. Letting e_i be the vectors that is all-zero except for a 1 at position i, we have

$$\begin{aligned} Mv_i &= \begin{pmatrix} \uparrow & & \uparrow \\ u_1 & \cdots & u_p \\ \downarrow & & \downarrow \end{pmatrix} \begin{pmatrix} \sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_p \end{pmatrix} \begin{pmatrix} \leftarrow & v_1 & \rightarrow \\ & \vdots & \\ \leftarrow & v_p & \rightarrow \end{pmatrix} v_i \\ &= U\Lambda V^T v_i \; = \; U\Lambda e_i \; = \; U\sigma_i e_i \; = \; \sigma_i u_i. \end{aligned}$$

   (b) Similarly,

$$M^T u_i \; = \; V\Lambda^T U^T u_i \; = \; V\Lambda e_i \; = \; V\sigma_i e_i \; = \; \sigma_i v_i.$$

   (c) Putting together the results of (a) and (b), we have

$$\begin{aligned} M^T M v_i &= M^T \sigma_i u_i = \sigma_i^2 v_i \\ M M^T u_i &= M \sigma_i v_i = \sigma_i^2 u_i \end{aligned}$$

   (d) The eigenvalues of MM^T are $\sigma_1^2, \sigma_2^2, \ldots, \sigma_p^2$ and the corresponding eigenvectors are $u_1, u_2, \ldots, u_p$.

   (e) The eigenvalues of $M^T M$ are identical to those of MM^T, but the eigenvectors are $v_1, v_2, \ldots, v_p$.

   (f) If M has rank k, the most significant k singular values are positive and the remaining $p - k$ singular values are zero.

2. *Matrix approximation*
   The rank-2 approximation of M is

$$\hat{M} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 4 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \end{pmatrix}$$

3. *Rank-1 matrices.*

   (a) The best rank-1 approximation of M is

$$\begin{aligned} \hat{M} &= \begin{pmatrix} -0.3863177 \\ -0.92236578 \end{pmatrix} \begin{pmatrix} 9.508032 \end{pmatrix} \begin{pmatrix} -0.42866713 & -0.56630692 & -0.7039467 \end{pmatrix} \\ &= \begin{pmatrix} 1.57454629 & 2.08011388 & 2.58568148 \\ 3.75936076 & 4.96644562 & 6.17353048 \end{pmatrix} \end{aligned}$$

   (b) It's generally possible to find a different pair of vectors. Let $a = \lambda u, b = \frac{1}{\lambda} v$, where $\lambda > 0$ and $\lambda \neq 1$. Then $ab^T = uv^T$.

   (c) $\widehat{M} = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_k u_k v_k^T$.