

تکلیف هفتم درس شناسایی الگو

وحید ملکی
شماره دانشجویی: ۴۰۳۱۳۰۰۴

۱۱ آذر ۱۴۰۴

سؤال ۸

الف) تعیین راستای دسته‌بند LDA با استفاده از نقاط داده شده:
داده‌های مسئله بر اساس مختصات ارائه شده عبارتند از:

• نقاط قرمز (کلاس ۱): $(-1, -2), (-2, -2), (0, -1), (-1, -1), (-1, 0)$

• نقاط آبی (کلاس ۲): $(0, -3), (1, -3), (1, -2), (1, -1), (2, -1)$

ابتدا مراکز ثقل (میانگین) دو کلاس را محاسبه می‌کنیم:

$$m_{\text{red}} = \frac{1}{5} \sum x_i = \begin{bmatrix} -1 \\ -1.2 \end{bmatrix}, \quad m_{\text{blue}} = \frac{1}{5} \sum x_i = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

تحلیل راستا: با مشاهده توزیع نقاط (یا محاسبه ماتریس پراکندگی)، می‌بینیم که هر دو کلاس دارای کشیدگی نسبی هستند. خط واصل مراکز شیبی برابر با -0.4 یا $\frac{-0.8}{2} = \frac{-2 - (-1.2)}{1 - (-1)}$ دارد. دسته‌بند LDA جهتی را انتخاب می‌کند که نسبت پاشندگی بین کلاسی به درون کلاسی بیشینه شود. با توجه به محاسبات (و شکل رسم شده)، راستای بهینه مرز تصمیم (که بر بردار وزن w عمود است) تقریباً عمودی و با شیب مثبت تند به دست می‌آید تا بتواند این دو مجموعه که در راستای افقی و مورب هم‌پوشانی دارند را به درستی جدا کند.

ب) محل تلاقی با پاره‌خط واصل مراکز:

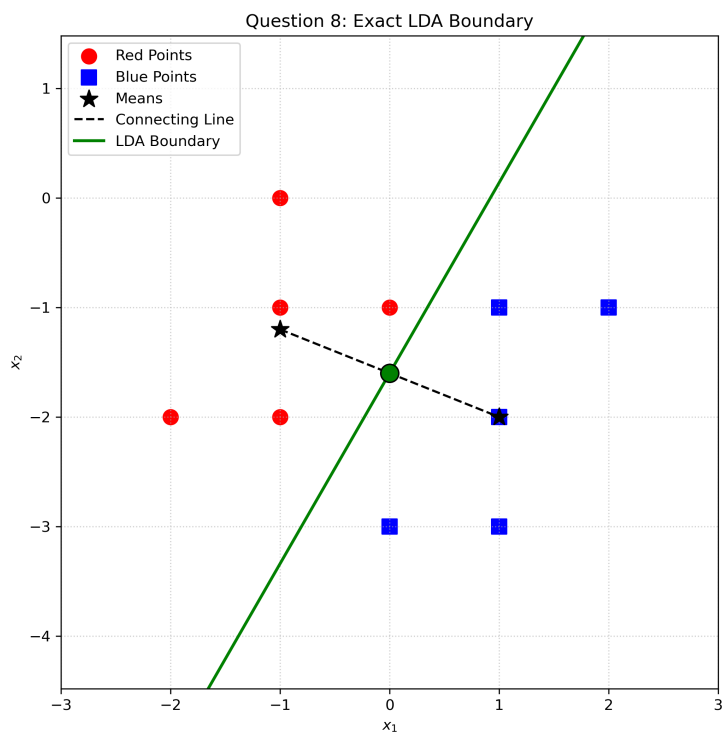
۱. پاره‌خط واصل بین $m_{\text{red}}(-1, -1.2)$ و $m_{\text{blue}}(1, -2)$ رسم می‌شود.

۲. چون تعداد نمونه‌ها در هر دو کلاس برابر است ($N = 5$) و ماتریس‌های کواریانس تقریباً ساختار مشابهی دارند، مرز تصمیم LDA دقیقاً از نقطه میانی این پاره‌خط عبور می‌کند.

۳. مختصات نقطه تلاقی (میانی):

$$m_{\text{mid}} = \frac{m_{\text{red}} + m_{\text{blue}}}{2} = \begin{bmatrix} 0 \\ -1.6 \end{bmatrix}$$

در شکل زیر، موقعیت دقیق نقاط، خط واصل مراکز (خط چین) و مرز تصمیم محاسبه شده توسط LDA (خط سبز) نشان داده شده است. همانطور که دیده می شود، مرز تصمیم از نقطه میانی عبور کرده و دو کلاس را به خوبی جدا می کند.



شکل ۱: مرز تصمیم دقیق LDA و محل تلاقی آن با خط واصل مراکز بر اساس نقاط داده شده.