Quadratic Programming

1. QP:

(a) \begin{align*}
\min_{x} & \quad q(x) = g^T x + \frac{1}{2} x^T H x \\
\text{subject to} & \quad Ax \geq b
\end{align*}

(b) $H$ is $n \times n$, real symmetric matrix.

c) $A = \begin{pmatrix}
a_1^T \\
\vdots \\
a_m^T
\end{pmatrix}$ is $m$ by $n$ real matrix.

d) $B = \begin{pmatrix}
b_1 \\
\vdots \\
b_m
\end{pmatrix}$

2. Convex problems:

(a) $H$ is positive semi-definite. ($x^T H x \geq 0$ for all $x$)

3. SVM:

(a) primal:

\begin{align*}
\min_{\alpha} & \quad \frac{1}{2} \alpha^T Q \alpha - e^T \alpha \\
\text{subject to} & \quad 0 \leq \alpha_i \leq C, i = 1, \ldots, l,
\end{align*}

(b) dual:

\begin{align*}
\min_{\alpha} & \quad \frac{1}{2} \alpha^T Q \alpha - e^T \alpha \\
\text{subject to} & \quad y^T \alpha = 0.
\end{align*}

4. KKT condition:
\( w(\alpha) = \sum_i \alpha_i y_i \phi(x_i) \) (4)

\( \bar{L} = \frac{1}{2} w(\alpha) \cdot w(\alpha) - \sum_i \alpha_i - \sum_i \delta_i \alpha_i + \sum_i \mu_i (\alpha_i - C) - \beta \sum_i \alpha_i y_i \) (5)

\( F_i = w(\alpha) \cdot \phi(x_i) - y_i \)
\( = \sum_j \alpha_j y_j k(x_i, x_j) - y_i \) (6)

\( \frac{\partial \bar{L}}{\partial \alpha_i} = (F_i - \beta) y_i - \delta_i + \mu_i = 0 \)
\( \delta_i \geq 0 \)
\( \delta_i \alpha_i = 0 \)
\( \mu_i \geq 0 \)
\( \mu_i (\alpha_i - C) = 0 \) (7)

(b) Simplified:
\( \alpha_i = 0 \Rightarrow y_i (F_i - \beta) \geq 0 \)
\( 0 < \alpha_i < C \Rightarrow y_i (F_i - \beta) = 0 \)
\( \alpha_i = C \Rightarrow y_i (F_i - \beta) \leq 0 \). (8)

5. Figure of convex problem in quadratic programming:

6. Different approaches:

(a) Chunking: divide large QP as several smaller QP problems by reducing the rows and columns corresponding to zero Lagrange multipliers.

(b) Osuna’s method: adding at least one example which violates the KKT conditions with maintaining constant size matrix for every QP sub-problem.

(c) Platt’s method: choose 2 lagrange multipliers to adapt analytically.
(d) Keerthi’s method: further analyze relationship between $F_i$ and $\beta$, and maintain 2 $\beta$s.

7. Discussion:

(a) Complement coding hidden in SMO method?

References


