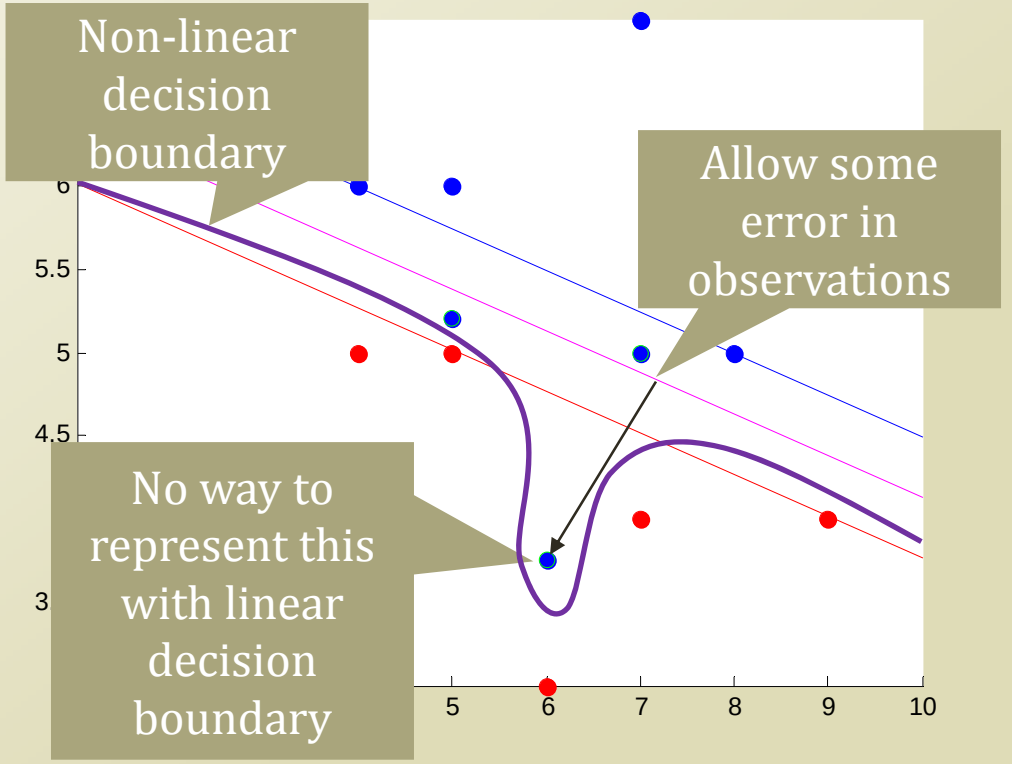
Support Vector Machine

Il-Chul Moon
Dept. of Industrial and Systems Engineering
KAIST
icmoon@kaist.ac.kr

SOFT MARGIN

“Error” Cases in SVM

- Data points that are
  - Impossible to classify with a linear decision boundary
- So called, “error” cases…
- How to manage these?
  - Option 1
    - Make decision boundary more complex
    - Go to non-linear
    - Any problem?
  - Option 2
    - Admit there will be an “error”
    - Represent the error in our problem formulation.
    - Try to reduce the error as well.
    - Any problem?



“Error” Handling in SVM

- How to handle
- Option 1)
  - Counting the error cases and reduce the counts
  - $min_{w,b}||w|| + C \times \#_{error}$
    $s.t.(wx_j + b)y_j \geq 1, \forall j$
  - Any problem?
- Option 2)
  - Introduce a slack variable
    - $\xi_j > 1$ when mis-classified
  - $min_{w,b}||w|| + C\sum_j \xi_j$
    $s.t.(wx_j + b)y_j \geq 1 - \xi_j, \forall j$
    $\xi_j \geq 0, \forall j$
  - Any problem?
- C = trade-off parameter