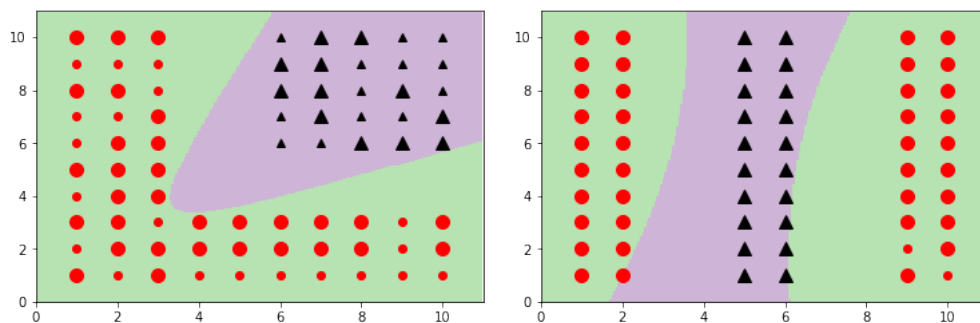
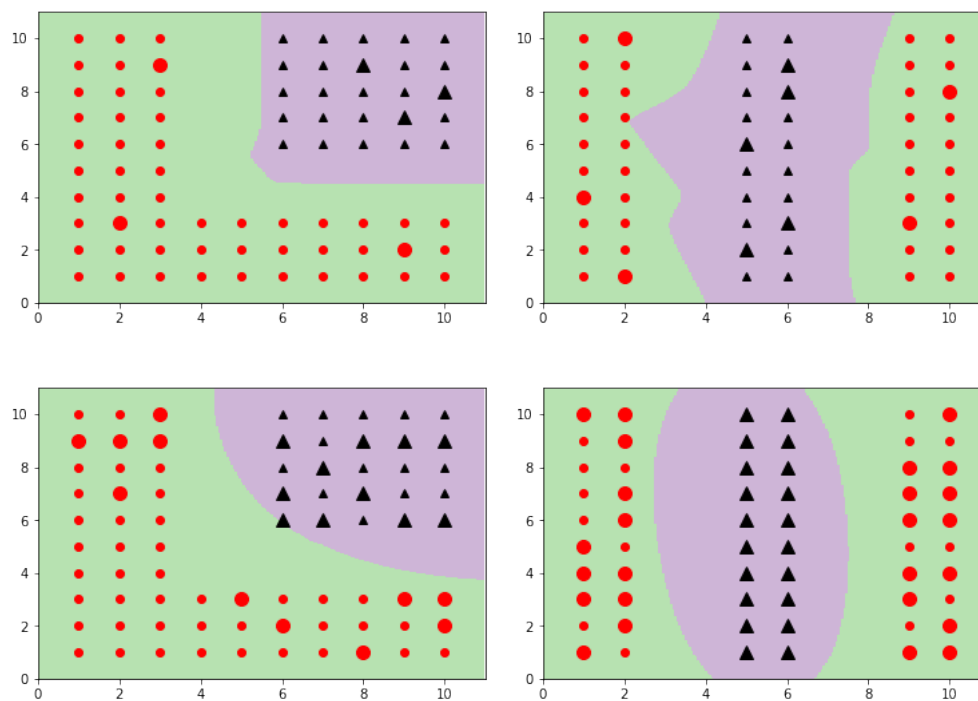


Week 8 — Solutions

1. $w = (1, 0, 0, -6, 3)$.
2. (a) There are four features of the form x_i , four of the form x_i^2 , and $\binom{4}{2}$ of the form $x_i x_j$, a total of 14.
(b) 14.
3. (a) Yes: the expanded representation is able to capture all quadratic boundaries, which include linear boundaries.
(b) Not necessarily: it might find a genuinely quadratic separator.
4. (a) α captures the number of updates on each point, so it is $(0, 2, 2, 1)$.
(b) Each update on point with label $+1$ increments b , while an update on a point with label -1 decrements b . Thus the final value of b is -1 .
5. (a) α has one coordinate per data point: a total of 36 dimensions.
(b) There are 6 support vectors, and thus 6 entries of α are > 0 .
(c) No entry of α is ever < 0 .
(d) Both sides of the boundary are the same size in the higher-dimensional space of the basis expansion, but the nonlinear mapping can make one side larger in lower dimension.
6. *Kernel Perceptron.*

Left: **data1**, Right: **data2**. The three rows correspond to the quadratic kernel, the RBF kernel with $\sigma = 1.0$, and the RBF kernel with $\sigma = 10.0$, respectively.





7. (a) The MNIST data is not linearly separable. Here are the training and test errors we got in a particular run; these might vary because the `sklearn` learning algorithm for SVMs often fails to converge.

C	train error (%)	test error (%)
0.01	14.94	15.44
0.1	10.92	11.62
1.0	10.97	11.90
10.0	11.51	12.02
100.0	11.80	12.88

- (b) Using a quadratic kernel with $C = 1.0$, we get training error 1.25% and test error 2.26%. The number of support vectors is 10133.