$X_{k+1} = AX_k + Bu_k + \omega_k$ $\quad$ O^i: obstacle centroid $i=1,\dots,\underset{\text{num obstacles}}{I}$

$\rightarrow$ $\quad X_k = [x_k, y_k]^T$

Robot dynamics LTI

MDP-state space

$s_k = [X_k, d_k]$, $d_k = f(X_k, O)$ => shortest distance to obstacle centroids

$s_k = g(x_k)$ $\quad$ $f(X_k) = \min_i \underbrace{\|X_k - O^i\|_2^2}_{f_i(X_k)} = \min_i f_i(x_k)$

$$f_i(x_k) = \|X_k - O^i\|_2^2 = (X_k - O^i)^T (X_k - O^i)$$

$$= X_k^T X_k - X_k^T O^i - O^{iT} X_k + O^{iT} O^i$$

$$\frac{\partial f_i}{\partial X_k} = 2X_k - 2O^i$$

$$\frac{\partial f}{\partial X_k} = 2X_k - 2O^j, \; j = \arg\min_i f_i(x)$$

Gradient of min()

$$g(x) = \min(g_1(x), g_2(x))$$

$$\frac{\partial g(x)}{\partial x} = \begin{cases} \frac{\partial g_1(x)}{\partial x}, & g_1(x) < g_2(x) \\ \frac{\partial g_2(x)}{\partial x}, & g_2(x) < g_1(x) \end{cases}$$

Dist. Robust Bellman Op.

$$\hat{\mathcal{J}} Q(s,a) = \inf_{P \in U} E_{s' \sim P}\Big[\underbrace{r(s,a,s') + \gamma \sum_{a' \in A} \pi(a'|s') Q(s',a')}_{h(s') = h(s,a,s')}\Big]$$

$h(s') = h(g(x'))$ => Re-define the $\hat{\mathcal{J}}$ operator around x' instead of s'

$$\hat{\mathcal{J}} Q(s,a) = \inf_{P \in U} E_{x' \sim P}[h(s,a,g(x'))] = \inf_{P \in U} E_{x' \sim P}[h(g(x'))]$$

=> Calculating the gradient of h

$$\frac{\partial h}{\partial x} = \frac{\partial h}{\partial s} \cdot \frac{\partial s}{\partial x} \qquad \frac{\partial s}{\partial x} = \underbrace{\left[1, \frac{\partial f}{\partial x}\right]}_{4}$$

$$\frac{\partial h}{\partial s} = \frac{\partial r}{\partial s} + \gamma \sum_{a' \in A} \left(\frac{\partial \pi}{\partial s} Q(s', a') + \pi(a'|s') \frac{\partial Q}{\partial s}(s', a') \right) \qquad \frac{\partial Q}{\partial s} \Rightarrow \text{Backprop}$$

$$\pi(a|s) = \frac{e^{Q(s,a)}}{\sum_{a' \in A} e^{Q(s,a')}}$$

$$\frac{\partial \pi}{\partial s} = \frac{\frac{\partial}{\partial s}\left(e^{Q(s,a)}\right)\left(\sum_{a' \in A} e^{Q(s,a')}\right) - e^{Q(s,a)}\left(\sum_{a' \in A} \frac{\partial}{\partial s}\left(e^{Q(s,a')}\right)\right)}{\left(\sum_{a' \in A} e^{Q(s,a')}\right)^2}$$

$$\frac{\partial}{\partial s}\left(e^{Q(s,a)}\right) = \underbrace{\frac{\partial}{\partial s} Q(s,a)}_{\text{backprop}} \cdot e^{Q(s,a)}$$

$\frac{\partial r}{\partial s}$ -> to be computed