Today
Lagrange Multipliers
KKT conditions
Augmented Lagrange

Announcements
Exam grades tomorrow
HW 7 due Thursday

Lagrange Multipliers

$$\mathcal{L}(\vec{x}, \vec{\lambda}, \vec{\sigma}) = f(\vec{x}) + \vec{\lambda} \cdot \vec{h}(\vec{x}) + \vec{\sigma} \cdot \vec{g}(\vec{x})$$

Lagrangian

Equality Lagrange Multipliers

Inequality L.M.

At the constrained optimum

$$\nabla_{\vec{x}} \mathcal{L}(\vec{x}^*, \vec{\lambda}^*, \vec{\sigma}^*) = \vec{0}$$

need to be the "right" values — the ones that make $\vec{x}^*$ satisfy the constraints

For the demo example

$$\mathcal{L}(x, \sigma) = x^2 + \sigma(1-x)$$

F.O.N.C.

Unconstrained

$$\nabla_{\vec{x}} f(\vec{x}) = \vec{0}$$

Constrained

$$\nabla_{\vec{x}} \mathcal{L}(\vec{x}^*, \vec{\lambda}^*, \vec{\sigma}^*) = \vec{0} \quad \text{stationarity}$$
$$\vec{h}(\vec{x}^*) = \vec{0}$$
$$\vec{g}(\vec{x}^*) \le \vec{0}$$
$$\vec{\sigma}^* \ge \vec{0}$$
$$\vec{\sigma}^* \circ \vec{g}(\vec{x}^*) = \vec{0}$$

Karush-Kuhn-Tucker KKT conditions

solving constrained optimization problems is mostly about solving KKT

$\circ$ elementwise multiplication, .* in matlab

"Complementary slackness"

Either $\sigma_i^* = 0$ or $g_i(\vec{x}^*) = 0$

Lagrange multipliers for inactive constraints are 0

How do we solve for $\vec{x}^*, \vec{\lambda}^*, \vec{\sigma}^*$ that satisfy KKT?

One method (of many): Augmented Lagrange

One equality constraint:

$$\hat{f}(\vec{x}; \lambda, \mu) = \underbrace{f(\vec{x}) + \lambda h(\vec{x})}_{\text{Lagrangian}} + \underbrace{\frac{\mu}{2} h(\vec{x})^2}_{\text{External Penalty}}$$

$\hat{f}$: Augmented Lagrangian

Augmented Lagrangian method, one equality constraint $h(\vec{x}) = 0$

input : starting point $\vec{x}_0$, initial multiplier $\lambda_0 = 0$, initial penalty $\mu_0 > 0$, penalty increase factor $\rho > 1$, tolerance E_s

output: $\vec{x}^*$, estimate of the constrained minimizer, and $f(\vec{x}^*)$

// augmented Lagrangian: Lagrangian plus quadratic penalty

$\hat{f}(\vec{x}; \lambda, \mu) = f(\vec{x}) + \lambda h(\vec{x}) + \frac{\mu}{2} h(\vec{x})^2$

$k \leftarrow 0$

while $|h(\vec{x}_k)| > E_s$ **do**

- $\vec{x}_k^* \leftarrow \underset{\vec{x}}{\text{minimize}}\ \hat{f}(\vec{x}; \lambda_k, \mu_k)$ // unconstrained solve started from $\vec{x}_k$ — Newton's / Gradient descent
- $\lambda_{k+1} \leftarrow \lambda_k + \mu_k h(\vec{x}_k^*)$ // update multiplier
- $\mu_{k+1} \leftarrow \rho \mu_k$ // increase penalty
- $\vec{x}_{k+1} \leftarrow \vec{x}_k^*$ // next starting point
- $k \leftarrow k + 1$

update

end

return $\vec{x}^* = \vec{x}_k$, $f(\vec{x}^*)$

when $h(\vec{x}) = 0$, $\hat{f}$ is just $\mathcal{L}$

will converge to $\lambda^*, \vec{x}^*$ when $h(\vec{x}) = 0$

For inequality constraints

$$\hat{f}(\vec{x}; \vec{\bar{\lambda}}, \mu) = f(\vec{x}) + \vec{\bar{\lambda}} \cdot \vec{\bar{g}}(\vec{x}; \vec{\bar{\lambda}}) + \frac{1}{2} \mu \left\| \vec{\bar{g}}(\vec{x}; \vec{\bar{\lambda}}) \right\|^2$$

$\vec{\bar{\lambda}}$: vector of Lagrange multipliers for equality and inequality

$f(\vec{x}) + \vec{\bar{\lambda}} \cdot \vec{\bar{g}}(\vec{x}; \vec{\bar{\lambda}})$: modified Lagrangian

$\frac{1}{2} \mu \left\| \vec{\bar{g}}(\vec{x}; \vec{\bar{\lambda}}) \right\|^2$: External penalty

$$\vec{\bar{g}}_j(\vec{x}; \vec{\bar{\lambda}}) = \begin{cases} h_j(\vec{x}) & \text{if constraint } j \text{ is an equality constraint} \\ g_j(\vec{x}) & \text{if } g_j(\vec{x}) \geq -\frac{\bar{\lambda}_j}{\mu} \\ -\frac{\bar{\lambda}_j}{\mu} & \text{otherwise} \end{cases}$$

Augmented Lagrangian penalty method

input : starting point $\vec{x}_0$, initial multipliers $\vec{\lambda}_0 = \vec{0}$, initial penalty $\mu_0 > 0$, penalty increase factor $\rho > 1$, tolerance E_s

output: $\vec{x}^*$, estimate of the constrained minimizer, and $f(\vec{x}^*)$

$$\bar{g}_j(\vec{x}) = \begin{cases} h_j(\vec{x}) & \text{equality constraint } h_j = 0 \\ g_j(\vec{x}) & \text{inequality } g_j \leq 0 \text{ with } g_j(\vec{x}) \geq -\lambda_j/\mu \\ -\lambda_j/\mu & \text{otherwise} \end{cases}$$

// augmented Lagrangian ($\vec{\lambda}$ contains multipliers for both $\vec{h}$ and $\vec{g}$)

$\hat{f}(\vec{x}; \vec{\lambda}, \mu) = f(\vec{x}) + \vec{\lambda}^{\mathsf{T}} \bar{g}(\vec{x}) + \frac{\mu}{2} \|\bar{g}(\vec{x})\|_2^2$

$k \leftarrow 0$

while $\|\bar{g}(\vec{x}_k)\| > E_s$ **do**

- $\vec{x}_k^* \leftarrow$ minimize $\hat{f}(\vec{x}; \vec{\lambda}_k, \mu_k)$ // unconstrained solve started from $\vec{x}_k$
- $\vec{\lambda}_{k+1} \leftarrow \vec{\lambda}_k + \mu_k \bar{g}(\vec{x}_k^*)$ // update multipliers
- $\mu_{k+1} \leftarrow \rho \mu_k$ // increase penalty
- $\vec{x}_{k+1} \leftarrow \vec{x}_k^*$ // next starting point
- $k \leftarrow k + 1$

end

return $\vec{x}^* = \vec{x}_k$, $f(\vec{x}^*)$

Will either push $g_j(\vec{x})$ to 0 OR $\vec{\lambda}_j$ to 0

When converged $\vec{\bar{g}}(\vec{x}^*; \vec{\bar{\lambda}}^*) \simeq \vec{0}$ so $\hat{f}$ is just $\mathcal{L}$