Convex Optimization
CMU-10725

Barrier Methods

Barnabás Póczos & Ryan Tibshirani

MACHINE LEARNING DEPARTMENT
Carnegie Mellon.
School of Computer Science

Barrier Methods (Interior methods)

Barrier Methods

$$\min_{x \in \mathcal{S}} f(x), \qquad (P)$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is continuous, and $\mathcal{S}$ is a constraint set in $\mathbb{R}^n$

Barrier program: replace (P) with the unconstrained problem (B(c)):

$$\min f(x) + \frac{1}{c} B(x)$$

where $c > 0$

Barrier function $B : \mathbb{R}^n \to \mathbb{R}$:

(i) B is continuous,

(ii) $B(x) \geq 0$ for all $x \in int(\mathcal{S})$

(iii) $B(x) \to \infty$ as $x \to \partial\mathcal{S}$

Barrier term: we don't let the algorithm leave S [Interior method]

Inequality Constraints

$$\min_{x \in \mathcal{S}} f(x) \qquad \mathcal{S} = \{x : g_i(x) \leq 0, i = 1, 2, \ldots, p\}$$

Example: $g_1(x) = x - b, \; g_2(x) = a - x$

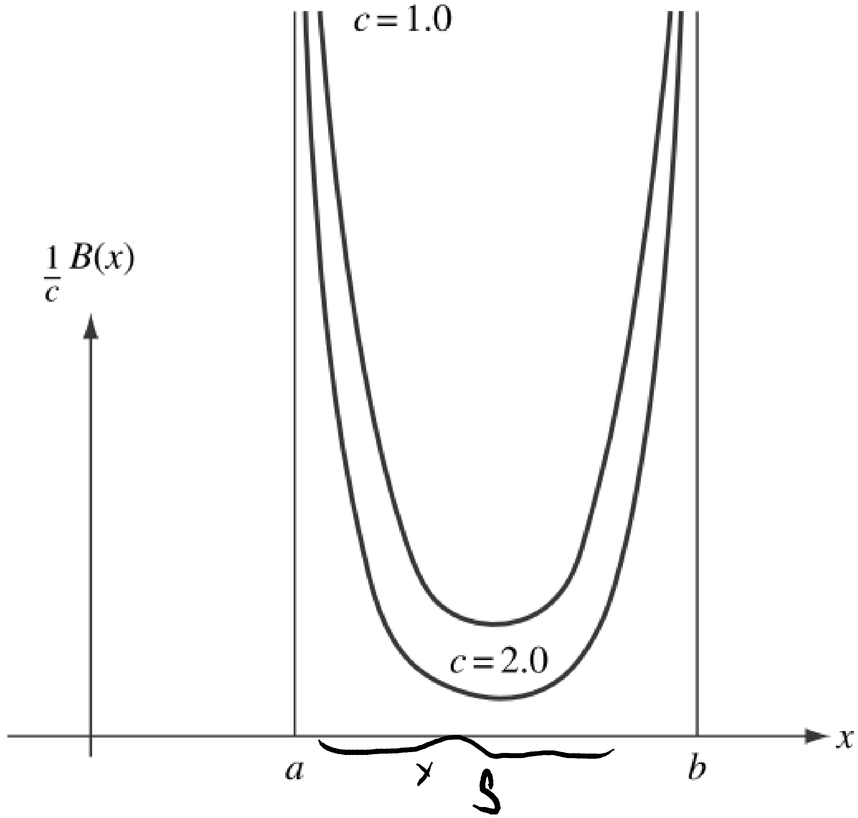
$a \leq x$
$x \leq b$

A useful barrier function in this case is:

$$B(x) = -\sum_{i=1}^{m} \frac{1}{g_i(x)}$$

Barrier program (B(c)):

$$\min f(x) + \frac{1}{c} B(x)$$



Logarithmic Barrier function

Problem: (P)

$$\min_{x \in \mathcal{S}} f(x) \qquad \mathcal{S} = \{x : \underline{g_i(x) \leq 0}, i = 1, 2, \ldots, m\}$$

Definition: Logarithmic barrier function

$$\underline{B(x) = -\sum_{i=1}^{m} \log(-g_i(x))}$$

Barrier method:

$$0 \leq c_1 < c_2 < \ldots < c_k < c_{k+1} \to \infty$$

$$B(c_k) \qquad x_k = \underset{\underline{x \in INT(S)}}{\operatorname{ARGMIN}} \; f(x) + \underbrace{\frac{1}{c_k}}_{\mu_k} B(x)$$

B(c) is still a constrained problem and looks more complicated than (P)

Advantage: B(c) can be solved with unconstrained optimization tools

Barrier Method

Penalty parameter:

$$0 < c_1 < c_2 < \ldots < c_k < c_{k+1} \to \infty$$

Barrier program:

$$B(c_k) \qquad x_k = \underset{x_k}{\mathrm{ARGMIN}} \; \underbrace{f(x_k) + \frac{1}{c_k} B(x_k)}_{r(c_k,\, x_k)}$$

Barrier Lemma:

1. $r(c_k, x_k) \geq r(c_{k+1}, x_{k+1})$
2. $B(x_k) \leq B(x_{k+1})$
3. $f(x_k) \geq f(x_{k+1})$
4. $f(x^*) \leq f(x_{k+1}) \leq f(x_k) \leq r(c_k, x_k)$

Proof of Barrier Lemma (1)

PROVE: $r(c_k, x_k) \geq r(c_{k+1}, x_{k+1})$

PROOF:

$r(c_k, x_k) \doteq f(x_k) + \frac{1}{c_k} B(x_k) \geq f(x_k) + \frac{1}{c_{k+1}} B(x_k)$

$c_k < c_{k+1}$

$\geq f(x_{k+1}) + \frac{1}{c_{k+1}} B(x_{k+1}) = r(c_{k+1}, x_{k+1})$

x_{k+1} IS OPTIMAL WITH c_{k+1}

Proof of Barrier Lemma (2)

PROVE $\quad B(x_k) \le B(x_{k+1})$

PROOF:

$$(*1) \quad f(x_k) + \frac{1}{c_k} B(x_k) \le f(x_{k+1}) + \frac{1}{c_k} B(x_{k+1})$$

x_k IS OPTIMAL WITH c_k

$$(*2) \quad f(x_{k+1}) + \frac{1}{c_{k+1}} B(x_{k+1}) \le f(x_k) + \frac{1}{c_{k+1}} B(x_k)$$

c_{k+1} IS OPTIMAL WITH x_k

$(*1) + (*2) \Rightarrow$

$$\frac{1}{c_k} B(x_k) + \frac{1}{c_{k+1}} B(x_{k+1}) \le \frac{1}{c_k} B(x_{k+1}) + \frac{1}{c_{k+1}} B(x_k)$$

$$\Rightarrow \underbrace{\left(\frac{1}{c_k} - \frac{1}{c_{k+1}}\right)}_{>0} B(x_k) \le \underbrace{\left(\frac{1}{c_k} - \frac{1}{c_{k+1}}\right)}_{>0} B(x_{k+1})$$

$$\Rightarrow B(x_k) \le B(x_{k+1}) \quad \square$$

Proof of Barrier Lemma (3)

PROVE $f(x_k) \geq f(x_{k+1})$

PROOF:

$$f(x_k) + \frac{1}{c_{k+1}} B(x_k) \geq f(x_{k+1}) + \frac{1}{c_{k+1}} B(x_{k+1})$$

x_{k+1} IS OPTIMAL WITH c_{k+1}

$$\geq f(x_{k+1}) + \frac{1}{c_{k+1}} B(x_k)$$

$B(x_{k+1}) \geq B(x_k)$ FROM LEMMA(2)

$$\Rightarrow f(x_k) \geq f(x_{k+1}) \quad \square$$

Proof of Barrier Lemma (4)

PROVE

$$f(x^*) \le f(x_{k+1}) \le f(x_k) \le r(c_k, x_k)$$

LEMMA (3)

PROOF:

$$f(x_*) \le f(x_k) \le f(x_k) + \underbrace{\frac{1}{c_k} B(x_k)}_{\ge 0} \doteq r(c_k, x_k) \qquad \square$$

x^* IS OPTIMAL IN S

$x_k \in S$

Primal Logarithmic Barrier Method for LP

Primal Log Barrier Method for LP

Primal problem

$$\min_{x} \; c^T x$$

$$\text{s.t.} \quad Ax = b \qquad \cancel{Ax \le b}$$

$$x \geq 0 \qquad x \in \mathbb{R}^n$$

Log barrier problem

$$f(x) = c^T x - \mu \sum_{j=1}^{n} \log(x_j)$$

$$B(\mu) \quad \min \; f(x)$$

$$\text{s.t.} \quad Ax = b$$

Derivatives:

$$g(x) \doteq \nabla f(x) = c - \mu D_x^{-1} e$$

$$e = (1, \ldots, 1)^T$$

$$G(x) = \nabla^2 f(x) = \mu D_x^{-2}$$

Primal Log Barrier Method for LP

Feasible solution of the log barrier problem

$\bar{x}$ FEASIBLE SOLUTION TO $B(\mu)$

$x^{+} = \bar{x} + \Delta x$

Quadratic approximation of f

$$f(x) = \underbrace{f(\bar{x}) + \underbrace{\nabla f(\bar{x})}_{g(\bar{x})}{}^{T} \overbrace{(x-\bar{x})}^{\Delta x} + (x-\bar{x}) \underbrace{\nabla^2 f(\bar{x})}_{G(\bar{x})} \overbrace{(x-\bar{x})}^{\Delta x} \cdot \frac{1}{2}}_{Q(\Delta x)}$$

Primal Log Barrier Method for LP

New quadratic problem:

$$\min_{\Delta x} \; Q(\Delta x)$$

$$\text{s.t.} \quad \left.\begin{aligned} A(\bar{x} + \Delta x) &= b \\ A\bar{x} &= b \end{aligned}\right\} \; \boxed{A \Delta x = 0}$$

Lagrange problem:

$$L(\Delta x, \pi_x) = g(\bar{x})^T \Delta x + \frac{1}{2} \Delta x^T G(\bar{x}) \Delta x - \pi_x^T [A \Delta x]$$

Primal Log Barrier Method for LP

Lagrange problem:

$$L(\Delta x, \pi_x) = g(\bar{x})^T \Delta x + \underbrace{\tfrac{1}{2}\Delta x^T G(\bar{x}) \Delta x}_{} - \underbrace{\pi_x^T [A \Delta x]}_{}$$

In the optimum:

$$0 = \frac{\partial L(\Delta x, \pi_x)}{\partial \Delta x} = \underbrace{g(\bar{x})}_{c - \mu D_x^{-1} e} + \underbrace{G(\bar{x})}_{\mu D_x^{-2}} \Delta x - A^T \pi_x$$

$$\begin{pmatrix} \mu D_x^{-2} & A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} -\Delta x \\ \pi_x \end{pmatrix} = \begin{pmatrix} c - \mu D_x^{-1} e \\ 0 \end{pmatrix}$$

$$x^+ = \bar{x} + \Delta x$$

Interior-point algorithm

MIN $f(x) = c^T x - \mu \sum_{j=1}^{n} \log(x_j)$

$x \geq 0$ s.t. $Ax = b$

MIN $c^T x$

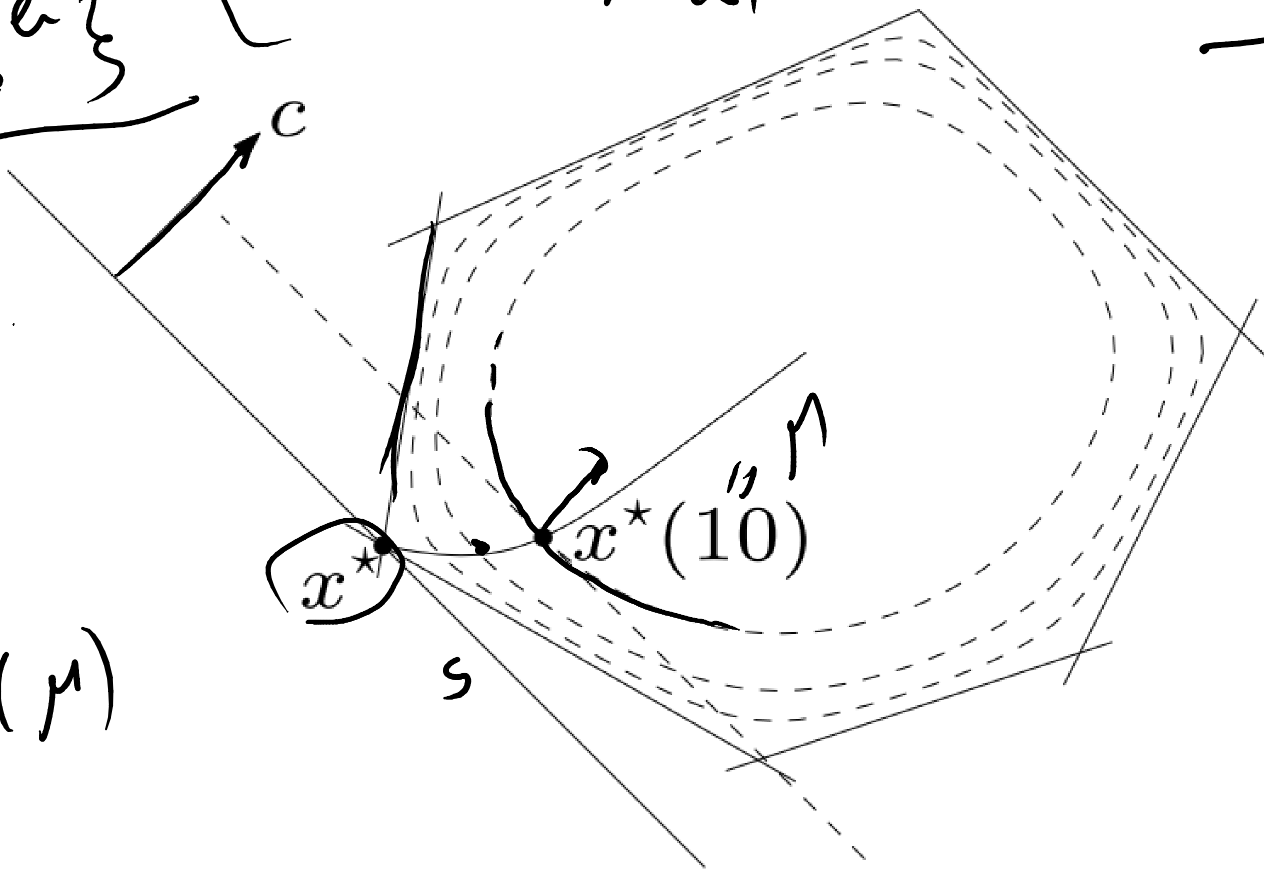
$\left.\begin{array}{l} Ax \leq b \\ x \geq 0 \end{array}\right\}$

$f(x) = c^T x - \overbrace{\mu \sum_{i=1}^{m} \log(b_i - a_i^T x)}^{\phi(x)}$

$n=2$

$m=6$

$x^*(\mu)$



Primal - Dual Logarithmic Barrier Method for LP

Symmetric Primal and Dual Forms

Symmetric form I [wikipedia]

(P) $\max \; c^T x$

s.t. $Ax \le b$

$x \ge 0$

(D) $\min\limits_{y} \; y^T b$

s.t. $y^T A \ge c^T$

$y \ge 0$

Duality gap:

$$\underset{\ge 0}{c^T x} \le (y^T A)\underset{\ge 0}{x} \le \underset{\ge 0}{y^T b}$$

Symmetric form II [Luenberger & Ye]

(P) $\min \; c^T x$

$Ax \ge b$

$x \ge 0$

(D) $\max\limits_{y} \; y^T b$

s.t. $y^T A \le c^T$

$y \ge 0$

Duality gap: $y^T b \le c^T x$

Asymmetric Primal and Dual Forms

Asymmetric form I [Jensen & Jonathan]

(P) $\max\ c^T x$

s.t. $Ax \le b$

$x \in \mathbb{R}^n$

$Ax + s = b$

(D) $\min_y\ y^T b$

s.t. $y^T A = c^T$

$y \ge 0$

Duality gap:

$b^T y - c^T x = (Ax+s)^T y - y^T A x = s^T y \ge 0$, $s \ge 0$

Asymmetric form II [Luenberger & Ye, R. Freund]

(P) $\min\ c^T x$

s.t. $Ax = b$

$x \ge 0$

(D) $\max_y\ y^T b$

s.t. $y^T A \le c^T$

$y \in \mathbb{R}^m$

$y^T A + s^T = c^T$

In what follows, we will use this form.

Duality gap: $c^T x - b^T y = s^T x \ge 0$

KKT conditions

Log barrier problem [We already know this]

$$\min_{x \in \mathbb{R}^n} f(x) = c^T x - \mu \sum_{j=1}^{n} \log(x_j)$$

$$\text{s.t. } Ax = b \qquad e = (1, \ldots, 1)^T$$
$$x \geq 0$$

Derivatives: $g(x) = \nabla f(x) = c - \mu D_x^{-1} e$

KKT stationarity condition of the log barrier problem

$$\underbrace{\nabla_x f(x)}_{} + y^T \underbrace{\nabla_x (b - Ax)}_{} = 0$$

$$[c - \mu D_x^{-1} e]^T + y^T(-A) = 0$$

$$c - \mu D_x^{-1} e = A^T y$$

Stationarity Condition Rewritten

The stationarity condition can be rewritten

$$c - \underbrace{\mu D_x^{-1} e}_{s} = A^T y$$

$$\mu D_x^{-1} e = s$$

$$\frac{1}{\mu} D_x \underbrace{s}_{D_s e} = e \quad \Rightarrow \quad \boxed{\frac{1}{\mu} D_x D_s e = e}$$

KKT Conditions

KKT Conditions

$$\left.\begin{array}{ll} Ax = b & x > 0 \\ A^T y + s = c & \\ \left[\ \frac{1}{\mu} D_x D_s e = e \right. & \\ \quad s \geq 0 & \end{array}\right\} \text{KKT}$$

Lemma

(x, y, s) SOLUTION OF KKT

$\Rightarrow$ • x IS FEASIBLE OF (P)

• (y, s) IS FEASIBLE FOR (D)

• $x^T s = e^T \underbrace{D_x D_s e}_{\mu e} = \mu \underbrace{e^T e}_{n} = \mu n$

Approximation of the KKT conditions

- $Ax = b \quad x > 0$
- $A^T y + s = c$
- $\frac{1}{\mu} D_x \overbrace{D_s e}^{s} = e$

$$\left\| \frac{1}{\mu} D_x s - e \right\| \le \beta$$

This is not linear in x and s!

$$\left\| \frac{1}{\mu} D_x s - e \right\| \le \beta$$

β-approximation of the stationarity condition:

Duality Gap of beta approximation

Lemma: [duality gap after β-approximation**]**

- $Ax = b,\ x > 0$
- $A^T y + s = c,$
- $\| \frac{1}{\mu} D_x s - e \| \leq \beta$

$\Big\}$ ($*$ KKT $-\beta$)

WE DON'T REQUIRE $s \geq 0$!

- IF $(\bar{x}, \bar{y}, \bar{s})$ IS A β-APPROXIMATE SOLUTION OF ($*$-KKT β)
- & $\beta < 1$

$\Rightarrow$ DUALITY GAP CAN BE LOWER & UPPER BOUNDED:

- $n\mu(1-\beta) \leq \underbrace{c^T \bar{x} - b^T \bar{y}}_{\text{DUALITY GAP}} \leq n\mu(1+\beta)$
- $\bar{x}$ IS FEASIBLE FOR (P), $(\bar{y}, \bar{s})$ IS FEASIBLE FOR (D)

Duality Gap of beta approximation

Proof:

PRIMAL FEASIBILITY: $\bar{x} \geq 0$ ✓

DUAL FEASIBILITY: $\bar{y} \in \mathbb{R}^m$ ✓

WE ONLY NEED TO SHOW THAT $\bar{s} \geq 0$

FROM $\left\| \frac{1}{\mu} D_x s - e \right\| \leq \beta$

$\Rightarrow -\beta \leq \frac{1}{\mu} x_j s_j - 1 \leq \beta$

$\Rightarrow 0 < (1-\beta)\mu \leq \underset{\overset{\vee\!|}{0}}{x_j} s_j \leq \mu(\beta+1)$

$\Rightarrow s_j > 0 \Rightarrow (\bar{y}, \bar{s})$ IS D FEASIBLE $\square$

DUALITY GAP:

$n\mu(1-\beta) \leq \sum_{j=1}^{n} x_j s_j = \underbrace{x^T s}_{\text{DUALITY GAP}} \leq n\mu(\beta+1)$

Primal-Dual Newton Step

Current iterate:

$\bar{x}, \bar{y}, \bar{s}$

- $\begin{cases} A\bar{x} = b, & \bar{x} > 0 \end{cases}$
- $\begin{cases} A^T\bar{y} + \bar{s} = c & \bar{s} > 0 \end{cases}$
- $\frac{1}{\mu} D_{\bar{x}} D_{\bar{s}} e - e = 0$

After update:

$\bar{x} \to \bar{x} + \Delta x$

$\bar{y} \to \bar{y} + \Delta y$

$\bar{s} \to \bar{s} + \Delta s$

$$\begin{cases} A(\bar{x} + \Delta x) = b \\ A^T(\bar{y} + \Delta y) + (\bar{s} + \Delta s) = c \\ \left(D_{\bar{x}} + D_{\Delta x}\right)\left(D_{\bar{s}} + D_{\Delta s}\right) e = \mu e \end{cases}$$

Newton steps:

$$A\,\Delta x = 0$$

$$A^T \Delta y + \Delta s = 0$$

$$D_{\bar{x}} \overbrace{D_{\Delta s} e}^{\Delta s} + \underbrace{D_{\Delta x} D_{\bar{s}} e}_{D_{\bar{s}} \Delta x} = \mu e - D_{\bar{x}} D_{\bar{s}} e - \cancel{D_{\Delta x} D_{\Delta s} e}$$

Primal-Dual Newton System

Primal-Dual Newton System

$$\left.\begin{array}{l} A\Delta x = 0 \\ A^T \Delta y + \Delta s = 0 \\ D_{\bar{s}} \Delta x + D_{\bar{x}} \Delta s = \mu e - D_{\bar{x}} D_{\bar{s}} e \end{array}\right\} \Rightarrow \begin{array}{l} \Delta x \\ \Delta y \\ \Delta s \end{array}$$

$$\mu_{k+1} = \alpha \mu_k$$

$$\mu_{k+1} = \frac{\bar{x}^T \bar{s}}{n} \cdot \frac{1}{10}$$

The Primal Dual Algorithm

STEP 0. INITIALIZATION

(x^0, y^0, s^0) FEASIBLE SOLUTION

STEP 1. $(\bar{x}, \bar{y}, \bar{s}) = (x^k, y^k, s^k)$

STEP 2. NEWTON STEP

$$\left.\begin{array}{l} \bullet\ A\Delta x = 0 \\ \bullet\ A^T\Delta y + \Delta s = 0 \\ \bullet\ D_{\bar{s}}\Delta x + D_{\bar{x}}\Delta s = \mu e - D_{\bar{x}} D_{\bar{s}} e \end{array}\right\} \Rightarrow (\Delta x, \Delta s, \Delta y)$$

STEP 3.

$x^{k+1} = \bar{x} + \Delta x$

$y^{k+1} = \bar{y} + \Delta y$

$s^{k+1} = \bar{s} + \Delta s$ $\quad 0 < \alpha < 1$

STEP 4 $\quad k = k+1. \quad \mu_{k+1} = \alpha \mu_k \quad \left[\text{OR } \mu_{k+1} = \frac{1}{10} \frac{\bar{x}^T \bar{s}}{n} \right]$

GO BACK TO 1 $\qquad$ CURRENT DUALITY GAP

Quadratic Convergence

Theorem [Local quadratic convergence of Primal-Dual Newton's method]:

- SUPPOSE THAT $(\bar{x}, \bar{y}, \bar{s})$ IS A β-APPROXIMATE SOLUTION OF $P(\mu)$ FOR SOME $0 < \beta < \frac{1}{2}$
- LET $(\Delta x, \Delta y, \Delta s)$ BE THE SOLUTION OF THE PRIMAL-DUAL NEWTON SYSTEM (* NEWTON)
- LET $(x', y', s') = (\bar{x}, \bar{y}, \bar{s}) + (\Delta x, \Delta y, \Delta s)$

$\Rightarrow$ (x', y', s') IS A $\frac{(1+\beta)}{(1-\beta)^2}\beta^2$-APPROX SOLUTION OF $P(\mu)$

Summary

- Penalty functions
- Barrier functions
- Primal log barrier method for LP
- Primal-Dual log barrier method for LP