15-830/630 – Problem Set #4

In this problem set, you will bring together several of the elements you have studied this far in class to solve the problem of scheduling generation and storage to minimize carbon emissions in a small smart grid setting. The basis for this problem is the IEEE 30 bus test case, which you saw in Problem Set 3, but instead of having six total generators producing a fixed amount of power (the slack bus and five PV buses), there are two generators, two energy storage facilities, and two wind farms. In addition, instead of just computing the solution for a single instant in time (given fixed demands), you'll need to use control methods to schedule generation and storage over time, taking into account varying load and different constraints of the generation. You'll combine this with forecasting methods to predict future wind and demand.

You can load the data for all these problems using the command

```
load ps4_data.mat
```

This will load several variables into MATLAB. For reference, all the loaded variables are listed below, but we'll also describe them in more detail in the relevant questions

- `gen = [1 2]` — bus indices of the controllable generators
- `storage = [5 8]` — bus indices of the storage facilities
- `wind = [11 13]` — bus indices of the wind farms
- `pq` — bus indices for all the PQ loads
- `winds` — matrix of time series data giving power production for the two wind farms over every hour in a year
- `loads` — matrix of time series data giving electrical demand for each of the PQ loads for every hour in a year (following convention, these values are negative to indicate consumed power)
- `Belec` — DC power flow approximation susceptance matrix; this matrix relates power and voltage angles via the DC power flow approximation $p = B\theta$. Note that for convenience we include the negative sign (that you saw in the derivation of power flow for the previous assignment) into B itself, so that you don't need to include it explicitly.
- `G,h` — line constraints; these specify constraints on the amount of power than can flow over different branches in the power network. The line constraints are given in the form $G\theta \leq h$.

Subsequent questions in this problem set will build upon previous ones, so be sure to start looking at these problems early!

1. **Autoregressive forecasting [35pts]** One of the chief challenges in smart grid planning is obtaining some estimate of future system conditions, be it predicted upcoming generation of a wind farm or predicted power consumption for a distribution station. One of the most common methods for forecasting time series data is the *autoregressive* model: given a time series

$$y_1, y_2, \ldots, y_T$$

the (deterministic) autoregressive model stipulates that the next value be some linear combination of k previous values

$$y_t = \sum_{i=1}^{k} \theta_i y_{t-i}.$$

(a) Show how we can represent the autoregressive model via a linear dynamical systems

$$\begin{aligned} x_{t+1} &= Ax_t \\ y_t &= Cx_t \end{aligned}.$$

Given an explicit definition of the components of the state vector x_t, and the A and C matrices. [Hint: make the state vector consist of a history of the y_t observations.]

(b) Given a sequence of observations

$$y_1, y_2, \ldots, y_T$$

show how we can learn the θ parameters using least squares. In other words, set up a least squares task to solve the optimization problem

$$\underset{\theta}{\text{minimize}} \sum_{t=p+1}^{T} \left(y_t - \sum_{i=1}^{p} \theta_i y_{t-i} \right)^2.$$

Write the solution to θ in terms of the normal equations

$$\theta = (\Phi^T \Phi)^{-1} \Phi^T z$$

for properly defined matrices Φ and vector z (we use z here instead of the normal y we used in the machine learning section to avoid confusion). Write out the explicit form of Φ and z in terms of the the y_t's.

(c) Given the learned θ parameters, and a sequence of past observed values $y_{t-p}, \ldots, y_{t-1}$, it is easy to apply the definition of the autoregressive model to forecast the next value y_t. How would you forecast the value after that? Give a formula for forecasting the next p values in the time series. Hint: this is easy if you explicitly form the A matrix mentioned in part (a), and think about powers of this matrix.

(d) Implement the above least-squares solution for the autoregressive model in MATLAB, as a function of the form

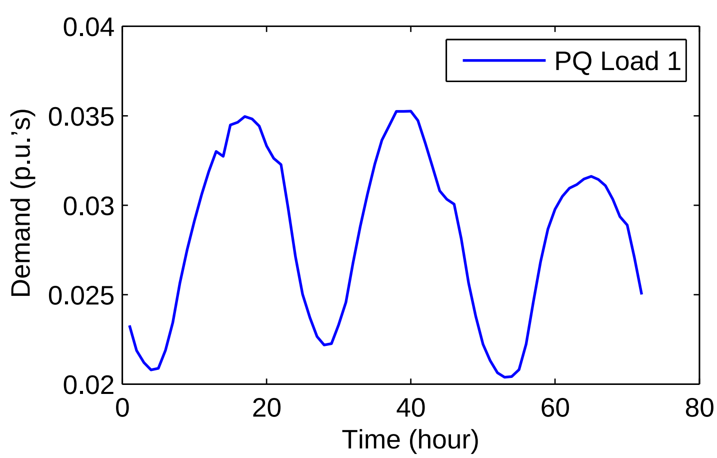
```
function theta = ar(y,p)
```

Hint: you might find the MATLAB functions `toeplitz` or `hankel` useful for this; using these functions, you can implement the above `ar` function in one line of MATLAB code (this isn't required for the problem, but it might just make the problem easier).

(e) The `winds` and `loads` variables (loaded from the `ps4_data.mat` above) contain hourly time series data for an entire year, given wind power (for the two wind turbine sites) and electrical demand (for each of the PQ nodes) respectively, where each time series is contained in a single column of the variable. Use the above `ar` function to build forecasting models for each of the time series, using $p = 24$ (i.e., you use a previous days worth of data to predict the next value).

For the first demand time series (`loads(1,:)`, this is the demand over the time frame $t = 4681, \ldots, 4752$ (three days in the summer):

```
plot(1:72,-loads(1,4681:4752))
```



Use your autoregressive model for this time series to forecast the demand for the time period $t = 4705, \ldots, 4728$ (the middle day) using the observations from the time period $t = 4681, \ldots, 4704$ (the first day). Plot the prediction on top of the actual demand.

2. **Generator and storage control [35pts]** In this problem, you will schedule generation and storage to minimize carbon emissions while respecting generator and power flow constraints. As mentioned above, we will be using a modified version of the IEEE 30 bus test. The grid has the following elements[1]

[1]Note that these precise generation and storage quantities do not necessarily represent realistic generators in terms of the absolute numbers, but they do capture the overall properties and relative strengths and weaknesses of the different types of generation and storage.

- There are two generators in the system, at buses 1 and 2. Bus 2 represents an open-cycle gas turbine (essentially a combustion system that directly burns the gas to move an electrical turbine); the advantage of such systems is that they can step up generation very quickly (in this setting, we assume it can change it's power generation arbitrarily quickly), but produce more carbon emissions than other approaches, due to their decreased efficiency. Bus 1 represents a combined-cycled gas turbine; these systems combine an open-cycle turbine with a steam generator powered by the exhaust heat of the turbine. The combined systems are about 60% more efficient that open-cycle systems, but because they rely on steam power (which has to be boiled before it can produce power), they have to ramp power up and down more slowly than the open-cycle generators.
- There are two energy storage facilities, a pumped-hydro facility at bus 5 and a battery storage facility at bus 8. As with generators, there are trade-offs involved: pumped hydro can store significantly more energy, but has lower efficiency than batteries; in this example, we assume both types of storage are roughly comparable in terms of their MW capacity.
- There are two wind farms in our system, at nodes 11 and 13. These wind farms generate power according to however the wind blows, and cannot be controlled to generate more power (though they can be turned off to stop generated power if needed). The power generated by these systems is given by the `winds` variables, discussed in the previous problem set.
- Finally, every node that is not one of the above is a PQ load, with power consumption given by the `loads` variable. Several of these nodes consume no power, which correspond to internal junction nodes in the network; although they consume no power, they are treated at PQ nodes because we know the real and reactive power consumption at the nodes (namely zero).

(a) We will define the state of the system as

$$x_t \in \mathbb{R}^4 = \begin{bmatrix} \text{Power from generator 1} \\ \text{Power from generator 2} \\ \text{Energy stored in pumped storage} \\ \text{Energy stored in battery} \end{bmatrix}$$

and the control input

$$u_t \in \mathbb{R}^6 = \begin{bmatrix} \text{Change in generator 1 power} \\ \text{Change in generator 2 power} \\ \text{Power used to charge pumped storage} \\ \text{Power used to charge battery} \\ \text{Power released by pumped storage} \\ \text{Power released by battery} \end{bmatrix}$$

Write the dynamics of the system in the form of a linear system

$$x_{t+1} = Ax_t + Bu_t$$

(i.e., explicitly write down A and B), capturing the fact that the pumped storage and battery storage are 80% and 90% efficiency respectively — i.e., if we put in 1 unit of power, this will only result in 0.8 or 0.9 units of energy stored (when we take power *out* of the storage, we assume that 100% of the energy is delivered).

(b) Write a set of constraints on the state and control variables, in the form

$$\underline{u} \leq u_t \leq \overline{u}, \quad \underline{x} \leq x_t \leq \overline{x}$$

that captures the constraints that

- The power produced by the generators, as well as the charge of the batteries, can never be negative.
- The pumped storage has a maximum energy capacity of 10 unit-hours, while the battery has a maximum capacity of 0.3 unit-hours.
- The combined-cycle generator can change its generator by at most 0.1 units per hour (we assume the open-cycle turbine can change as much as desired between hours).
- The pumped storage and battery can consume or produce at most 0.3 units of power at each hour.

Make sure that the constraints you put in place enforce the fact that the storage must lose energy: i.e., putting in 1.0 units of power will increase the pumped storage by 0.8 unit-hours, but you certainly can't take out 0.8 unit hours to get 1 unit of power.

(c) Let $p_t \in \mathbb{R}^{30}$ denote the power injections at each bus at time t. Define the elements p_t^{gen}, p_t^{wind}, and p_t^{storage} in terms of the state and control variables, and the variable $\mathrm{wind}_t \in \mathbb{R}^2$, the wind produced by the wind farms (where we can use all the power generated by the turbines if desired, but could also use less). The power consumed by the 24 PQ loads at time t is assumed to be given by the variables $\mathrm{loads}_t \in \mathbb{R}^{24}$, and cannot be adjusted.

(d) Combining the parts above, write down the model predictive control task as a linear optimization problem over the variables $x_{1:T}$, $u_{1:T}$, $p_{1:T}$, $\theta_{1:T}$. The objective of the optimization problem is to minimize carbon emissions, which are equal to the sum of $(x_t)_1 + 2(x_t)_2$ over all time (the power of the combined-cycled generator plus two times the power of the open-cycle generator, to take into account the fact that it is less efficient). You will want to include all the constraints mentioned above, plus the fact that x_1 equals some constant x^{init} at all indices *except* the second generator[2] (this accounts for the fact that the second generators power can be adjusted arbitrarily, and will need to be adjusted at the first state to ensure the total generation equals the total demand), and also the power flow constraints

$$p_t = B^{\mathrm{elec}}\theta_t, \quad (\theta_t)_1 = 0, \quad G\theta_t \leq h. \ \forall t$$

[2]**Important: The initial version of this problem set used the convention x_0 for the initial state, but this caused some confusion, since the dynamics would have to be adjusted slightly from their above form. Instead, forget everything from the initial version about the x_0 state, and just optimize over $x_{1:T}$ as normal, but also be sure *not* to constrain the second element $(x_1)_2$ to be equal to x_2^{init} (i.e., just enforce the constraint for the indices 1,3, and 4).**

where B^{elec} is the DC power flow susceptance matrix (denoted with the superscript "elec" just to avoid confusing with the "B" matrix from the dynamics) and where G and h enforce line flow constraints (in our case, the fact that no line can transmit than 1 unit of power). You can write the objective and constraints in whatever form is easiest (you don't need to put the problem into standard form).

(e) Implement the above optimization problem as a MATLAB function of the form

```
[f,X,U,P,Theta] = optimize_power(x0, wind, loads, Belec, G, h, ...
                                 gen, storage, wind, pq)
```

where `x0` is the initial state of the system; `winds` and `loads` are respectively a $2 \times T$ and matrix of future wind powers and a $24 \times T$ matrix of future loads (either true or forecasted in both cases); `Belec`, `G`, and `h` are matrices that capture the power flow constraints mentioned above; and `gen`, `storage`, `wind`, and `pq` are the indices of the different bus types. You should infer the length of the horizon T by inspecting the size of these input variables. For the output variables, `f` should be a scalar value indicating the objective function at the solution; `X` should be a $4 \times T$ dimensional matrix of all the states; `U` a $6 \times T$ matrix of all chosen controls; `P` a $30 \times T$ matrix of all power injections; and `Theta` a $30 \times T$ variable of all voltage angles in the DC power flow approximation.

You should use YALMIP to solve this optimization problem, and you'll find the problem much easier to solve if you define all the output variables (other than `f`) directly as matrix variables, i.e.,

```
U = sdpvar(6,T);
```

and write as many constraints as possible in matrix form.

(f) Use the function above to solve two different cases:

i. Solve the control problem using the real future values of power consumption and wind, starting at time $t = 4681$ with a time horizon $T = 24$ and an initial state of $x = (0, 0, 0, 0)$, i.e., using the call

```
[f,X,U,P,Theta] = optimize_power(zeros(4,1), winds(:,4681:4740), ...
                                 loads(:,4681:4704), Belec, G, h, ...
                                 gen, storage, wind, pq)
```

ii. Solve the problem using the same settings as above, but with the forecasted wind and load data (over the horizon $T = 24$ from your auto-regressive model in Question 1).

In both cases, report the resulting states and controls. Qualitatively describe the actions taken by this controller (e.g., how does the solution trade off between the two types of generation, when does it use the battery, etc). Does using the forecasted values change the behavior of the controller?

3. **(15-830 only) Model predictive control [30pts]**

(a) Because we are continually solving optimization problems many times, model predictive control is one settings where we often really do want very fast solutions to the optimization problems. With this as our motivation, rewrite the above optimization problem and MATLAB function as a linear program in standard form (this is actually a bit different from the standard form, since we explicitly include upper and lower bounds for the variables, but most solvers can directly include these constraints without having to write them in terms of the remaining inequality constraints)

$$\begin{aligned} \underset{z}{\text{minimize}} \quad & c^T z \\ \text{subject to} \quad & Az \leq b \\ & A_{\text{eq}} z = b_{\text{eq}} \\ & l \leq z \leq u \end{aligned}$$

where z is the optimization variable, and c, A, b, A_{eq}, b_{eq}, l, and u are optimization variables. You can then solve the problem using, for example, the CPLEX LP solver:

```
[z,f] = cplexlp(c, A, b, Aeq, beq, l, u);
```

Verify that your solution gives the same answer as the part you developed in question 2. It is important to use *sparse* matrices for A and A_{eq}: you can construct these using the MATLAB commands `sparse`, `speye`, and `spdiags` (see the MATLAB help page for the documentation for these functions).

(b) Equipped with your fast solver, run model predictive control to determine how to schedule power for the period $t = 3505, \ldots, 5520$, starting from an initial state $x = 0$. Recall that the procedure for applying MPC is as follows:

For $t = 3505, \ldots, 5520$:

- Use the previous $t - 24, \ldots, t - 1$ time steps to forecast wind and load (using the autoregressive model) at each bus for the upcoming $t, \ldots, t + 23$ time steps.
- Solve the control optimization problem (the LP formulation) to get a series of desired controls and states.
- Execute the first returned control action, u_1.

Plot a graph of the power produced or consumed by the generators, storage, and wind farms over the period. Also report the final objective function,that is:

$$f = \sum_{t=3505}^{5520} ((x_t)_1 + 2(x_t)_2)$$